UNIVERSITY OF WARWICK
WARWICK BUSINESS SCHOOL

International Diffusion of New Technology

Thesis submitted to the University of Warwick
for the degree of

Doctor of Philosophy

By

Anni-Maria Katariina Pulkki-Brännström

September 2009

Contents

Acknowledgements

I would like to dedicate this work to Niklas Brännström on whose encouragement, support and love I have come to depend. *Vi har båda gjort två av de här nu, så gör det fyra tillsammans?*

I would like to thank Professor Paul Stoneman of Warwick Business School, and Otto Toivanen, director of the Helsinki Centre for Economics Research, for excellent supervision.

Financial support from the Economic and Social Research Council UK and the Social Insurance Institution of Finland (KELA) is gratefully acknowledged.

My family's support has been essential in all ways and I am very grateful for it. I would like to take this opportunity to remember the late Helmi Laitinen, Sinikka Pulkki and Ulla-Maija Voipio who showed an example to later generations with their dedication to learning and teaching. For support particularly in the last one and a half years I thank Wu Yi and Pam Payne for their friendship, and BBC Radio Three, BBC Radio Four, KLM Royal Dutch Airlines, and Skype™ for valuable services.

Declaration

To the best of my knowledge and unless otherwise stated, all the work in this thesis is original.

No part of this thesis has been submitted for a degree at another university.

Chapter 2 has been jointly written with Paul Stoneman. All other work is entirely my own. None of this thesis has been previously published anywhere.

The econometric modelling software that has been used in this work is PcGive.

Abstract

This study explores the international diffusion of new technology i.e. changes over time in the extent to which world output is produced using, or world consumption is made up of products incorporating, specific new technologies. This topic has received relatively little attention in the literature. Many of the theoretical arguments developed in the literature for the study of domestic diffusion are here systematically applied to international diffusion for the first time.

We propose that patterns of international diffusion derive from two related processes: inter-country diffusion or the extensive margin, and intra-country diffusion or the intensive margin. We start with a study of the relative importance of these two processes. Using data on four technologies we show that the relative importance of the intensive to the extensive margin increases over time. The same pattern was identified by Battisti and Stoneman (2003) in their study of the importance of inter- and intra-firm diffusion in domestic diffusion.

The main body of the thesis is concerned with the question how (if at all) does international diffusion affect domestic diffusion? Two theoretical arguments are explored: the first uses an epidemic and the second a decision-theoretic model. The models are extensions of the seminal models of Bass (1969), Mansfield (1961) and Reinganum (1981b). Two specific hypotheses arise, namely that international diffusion affects domestic diffusion through: i) an exogenous learning effect or inter-

country spillovers; and ii) a negative stock effect. The hypotheses have contradictory empirical implications.

The epidemic model is tested using data on steam- and motor ship diffusion. We find evidence of spillovers however the direction of the effect is not robust across countries. We discuss the time-series properties of the data, which is rarely done in the literature, and find some problems which may partly explain the results.

We then develop an international stock effect hypothesis using a decision-theoretic model based on the closed economy model of Reinganum (1981b). This allows for firm heterogeneity in production costs. We discuss how heterogeneity impacts on international diffusion patterns when some of that heterogeneity is on the country-level.

Empirically we find evidence of an international stock effect in the diffusion of the basic oxygen furnace. A number of explanatory variables which capture cross-country differences in production and adoption costs are also significant.

List of Figures

List of Tables

1 Introduction

1.1 Motivation and concepts

Technological change is the driving force of economic growth. Schumpeter (1934) identified three phases of technological change: invention or the conception of a new idea, innovation which is the first use or marketing of a new product or process that embodies that idea, and diffusion which is the spread of the use of that product or process over time across its potential market. It is invention and innovation that continue to receive the most attention in research and society at large. However, it is only through the use or diffusion of new technologies that their benefits arise (Stoneman, 2002) and as such although relatively ignored, diffusion is of key importance in the growth of economic welfare. In fact, as for most countries most new technologies are invented and first used elsewhere, the extent to which new technologies invented elsewhere are adopted and used, or diffused, may be the key determinant of such growth.

Diffusion refers to the development of the usage of new technology by households (new products) or firms (new processes) across and within different industries and countries. Most literature takes a closed economy or single economy viewpoint in analysing the diffusion of new technology (see below). Here however we take an international perspective on diffusion. The international perspective relates to how and why a new technology is used to an increasing extent across the world, rather than an individual country, over time. It thus concerns, for example, the proportion of world output produced using a new technology (e.g. basic oxygen steel, or cargo kilometres registered using airplanes), or world consumption made up of products incorporating a new technology (e.g. digital TVs).

At the industry level, Battisti and Stoneman (2003) separate out the effects of inter- and intra-firm diffusion on overall industry diffusion. They find that at the early stages of diffusion inter-firm diffusion (the spread across firms) contributes more to growth in overall use than intra-firm diffusion (extent of use by each firm), while at later stages intra-firm diffusion is more important. They argue that to understand diffusion, one needs to understand both processes. Pursuing a similar approach, and as also suggested by Comin et al. (2006), at the international level, diffusion as defined above can be seen to derive from two processes: i) technology spreading across countries (inter-country diffusion); ii) technology spreading within countries (intra-country diffusion). We call the two resultant margins the extensive and intensive margins respectively.

This analytical approach that distinguishes between the intensive and extensive margins of diffusion is still new, and the present study contributes to that discussion. However, in addition to just "accounting" for diffusion as the result of changes in the intensive and extensive margin, the two processes although previously independently studied (in a variety of different literatures) have not until recently been analysed as two components of an integrated process. We are particularly interested in the dynamics of and possible interactions between inter- and intra-country diffusion in order to provide a more detailed understanding of international diffusion. The three main related objectives of this study are thus to: i) illustrate how new empirical insights can be gained from examining diffusion on the international level; ii) make a theoretical case for a link between national and international diffusion; iii) provide empirical support for this link.

1.2 Overview

This study uses the analytical techniques of diffusion research to study international diffusion, a topic which has received relatively little attention in the literature. The research contributes both to the theoretical and empirical diffusion literatures.

Our analytical approach is to distinguish between the intensive and extensive margins of diffusion so that international diffusion consists of two processes: diffusion within and diffusion across countries. We begin with a study of the relative importance of the two margins in Chapter 2. Using data on four technologies from HCCTAD (see below), we identify the same patterns on the international level as Battisti and Stoneman (2003) found in their study of domestic inter- and intra-firm diffusion of new technologies. The technologies and units of measurement are postal services (units of mail handled per $GDP); electricity (MWhrs of output per $GDP); telephones (number of mainland lines per $GDP); and basic oxygen steel (proportion of crude steel output).

The main body of the thesis is then devoted to the two questions that follow from this finding. The first question is, how (if at all) does international diffusion affect domestic diffusion? The hypothesis is that it does, and specifically that the extent of use elsewhere affects intra-country diffusion. We call this the international effect or the inter-country spillover hypothesis. From this arises the second and closely related question: how does the extent of use in any one country impact upon diffusion in other countries?

To answer these questions many of the theoretical arguments in the diffusion literature are here systematically applied to international diffusion for the first time.

In particular, the decision-theoretic framework appears well suited to the analysis of international diffusion even if it has rarely been applied to country-level data. The aim is not to present a single model or framework for the analysis of international diffusion, or a single model of how the intensive and extensive margins are linked. Therefore, we do not construct an encompassing model but rather our aim is to make a case for the international effect within two of the main theoretical frameworks. In the future it may be possible to formulate a general model of international diffusion but for now examining the determinants and interactions of inter- and intra-country diffusion in this way is a considerable step towards a more detailed understanding of international diffusion.

Specifically we explore two models, one using an epidemic and the other using a decision-theoretic foundation. The models are extensions of the seminal models of Bass (1969), Mansfield (1961) and Reinganum (1981a,b) and the two specific hypotheses that arise are that international diffusion affects domestic diffusion through: i) an exogenous learning effect; ii) a stock effect. The hypotheses are contradictory in that learning implies a positive relationship while the stock effect implies a negative effect (see below). We present the theoretical arguments in Chapters 3 and 4 and the hypotheses are tested using data on two technologies, steam- and motor ships and basic oxygen steel, in Chapters 3 and 5 respectively.

Our empirical contributions are to show the changing relationship between the intensive and extensive margins of diffusion, and support for the hypothesis that international diffusion affects domestic diffusion throughout the diffusion process. Our findings regarding other determinants of diffusion – drawing on arguments based in endogenous growth models – also contribute to that discussion. Although we use specific technologies as examples, our primary concern is with the determinants of diffusion on a more general level. This means that the empirical models may seem simplistic to readers with a special interest in the technologies under study; and that we put more value on parsimony of the model over fit.

1.3 Diffusion literature and studies of international diffusion

This thesis does not include a separate chapter reviewing the relevant literature. The reasons for this are twofold. Firstly, individual chapters include their own short literature reviews, and secondly there are already many comprehensive surveys of the literature – see Stoneman (2002), and also Geroski (2000) and Hall (2004) – and further effort in this direction would seem to offer minimal returns. Here, we present a short review of the theoretical and empirical issues that are most relevant to the present study. Relevant theoretical ideas on international diffusion are dispersed and empirical studies are few. We review also how analytical techniques of diffusion studies have been applied to international diffusion and discuss some of the macroeconomic literature that discusses international diffusion.

This study belongs to the literature on the economics of diffusion. The two main theoretical approaches to the study of diffusion are the epidemic (disequilibrium) approach, in which the role of information is emphasised, and the decision-theoretic (equilibrium) approach which focuses on the costs and benefits of adoption. Broadly speaking, epidemic models are more widely used across various disciplines while decision-theoretic models are found particularly in the economics literature. Karshenas and Stoneman (1993) summarise the determinants of diffusion in the various models as rank, stock, order, and epidemic effects. In this study we develop our analysis of international diffusion particularly using the epidemic framework (Chapter 3) and the stock effect hypothesis (Chapters 4 and 5).

1.3.1 Studies of international diffusion – a brief review

The international dimension features little in diffusion literatures. The empirical studies that exist can be divided into studies of: i) the diffusion of a single technology in a country over time; ii) differences in diffusion across countries (typically in a cross-section); and iii) spillover studies. We begin with the last category as this can be situated in the macroeconomic literature rather than diffusion studies and follow by looking at the other two.

Macroeconomic literature views technology as a public good that contributes to the nation's stock of productive potential. This is based on a view of technology as public knowledge, which the innovator creates and which can then be used by others to create more knowledge (Barro and Sala-i-Martin 1997, Eaton and Kortum 1999). This leads to models where diffusion is the imitation of technologies innovated abroad, which is a very different approach from the diffusion literature per se. Imitation increases productivity through an increase in the variety or quality of intermediate inputs (Barro and Sala-i-Martin 1997) or of final goods (Grossman and Helpman 1993). Our concern is that this literature takes a too simplistic view on diffusion, as the focus is on the "stock" of technological knowledge rather than the extent to which technologies are used.[1] The length of time that diffusion takes is

[1] Also as Stoneman (2002:5) points out information is not a public good because it is excludable to a degree through secrecy, patenting and copyright.

typically not recognised as the benefits of technology are assumed to be immediately realised. That is, the intensive margin of international diffusion is ignored.

Macroeconomic literature contributes however by analysing the uncertainty faced by producers of new technology competing in an international market. Grossman and Helpman (1993) argue that after innovation production will shift to wherever production costs are the lowest and this contributes to diffusion. In their model Northern innovators expect Southern firms to copy their products and attract all customers with a lower price. Southern firms expect Northern firms to innovate a superior quality product similarly capturing the market. This uncertainty about future profits lowers the incentives to innovate and imitate.

Spillover studies examine the positive externalities that occur when technology is transferred across borders. These studies share a certain focus on information-spreading as in epidemic models, but otherwise the literature is closely related to macroeconomic studies of growth and productivity. A survey is provided by Keller (2004). Trade and foreign direct investment (FDI) are seen as the channels of diffusion because imports and exports facilitate the transfer of embodied knowledge while FDI and other personal contact transfers tacit knowledge. In spillover regressions productivity is regressed on R&D, patents, or FDI and a positive relationship is expected. Keller (2004:779) concludes that there is lack of agreement in the literature on the quantitative effects but that "the international dimension of technological change is of key importance for most countries… the ongoing interaction with foreign firms and consumers seems to be a process of knowledge discovery for firms that cannot be had from interacting only with other domestic firms."

In this literature, diffusion within the country is discussed in terms of firm-to-firm spillovers. Branstetter (2001) for example studies patent data and R&D expenditures and finds evidence of positive domestic spillovers but foreign spillovers are not statistically significant (although they are significantly different from the domestic effect). The author suggests that internationally, a rivalry effect may dominate any spillovers so that firms do not benefit from research done in another country but they do benefit from research by other domestic firms. Interestingly, using a very different model we find a similar result in Chapter 5 namely that the domestic extent of use is positively related to further use while foreign diffusion is negatively related.

A few studies have shown that intra-country diffusion is faster in countries which lag behind in terms of inter-country diffusion (date of first adoption). However, these studies fail to link intra-country diffusion with the extent of use in other countries. Dekimpe et al. (2000) and Perkins and Neumayer (2005) use a spillover argument to analyse how the date at which intra-country diffusion is completed ("penetration" or "confirmation") depends on the date of first adoption in that country. The theoretical explanation given is that latecomers enjoy some advantage either through a reduction in uncertainty about the technology (Dekimpe et al.) or that the net benefits of use increase through improvements in the technology and reduction in the costs of adoption (Perkins and Neumayer). In effect, adoption by others increases the pool of knowledge about the technology and that knowledge is a public good. Dekimpe et al. (2000) reach their result by studying three different hazards: from first to "full" adoption (i.e. start and end point of intra-country diffusion); no use to first use; and no use to full adoption.[2] They find that the passage of time has a positive effect on each hazard.[3] Perkins and Neumayer (2005) present a hazard model of the diffusion of three technologies in OECD and developing countries where the event of interest is "penetration", defined as the completion of intra-country diffusion. They link the latecomer advantage hypothesis with convergence and find evidence that developing countries benefit of a latecomer advantage: the date of first use has a positive effect on the speed of intra-country diffusion, that is, the time between first use and the "penetration" level is shorter if first adoption occurs later in time.

The problem with both of these models is that they still fail to fully appreciate the two dimensions of international diffusion, the intensive and extensive margins. The passage of time, rather than the extent of international diffusion, is the key variable empirically which is confusing because the causal effect is attributed to spillovers from learning-by-doing rather than time itself.[4] It seems to us that a model is needed where time and international diffusion are both given a clear theoretical basis in a model of intra-country diffusion in order to answer the question why and how if at all, does international diffusion matter? Perhaps the extension suggested by Dekimpe et al. (2000) can establish such a clear link between inter- and intra-country diffusion; they suggest making the probability of full adoption a function of the probability of first adoption (rate dependence rather than state dependence).

---

[2] The hazard rate is the probability of an event occurring, given that it has not yet occurred. Dekimpe et al. (2000) argue that the "events" of first adoption and full adoption are closely interlinked and potentially influencing one another, however managerially the decisions are separate problems.

[3] A Weibull baseline hazard is used and in addition time since first domestic use is used in the equation for first use to full use.

[4] Perkins and Neumayer (2005) also include a regional diffusion variable (with the world divided into eight geographical regions). They explicitly do not want this to reflect the passage of time but local information spillovers (learning through contact with users or suppliers) which is why regional diffusion is measured relative to the global average.

However, we are not aware of any study that has proceeded with this suggested extension. Tellis et al. (2003) are closer to appreciating the full diffusion process. They include a time-varying measure of the extent of use elsewhere ("number of prior takeoffs") in a model where the dependent variable is the time to "takeoff" – a level of intra-country diffusion between first use and the end point. However here again international diffusion is taken to affect one particular point in domestic diffusion, not the whole process of intra-country diffusion.

The link between international and domestic diffusion is analysed even less in the numerous empirical studies that compare intra-country diffusion paths of a particular technology in a sample of countries. For example, Caselli and Coleman II (2001) study computer adoption in a model which is influenced by the spillover literature. They find that a country's openness to (manufacturing) imports has a positive effect on computer investment (imports), which is interpreted as evidence of knowledge spillovers. Comin et al. (2006) analyse a number of issues related to international diffusion such as the shape of intra-country diffusion curves, whether the speed is generally increasing, and whether there are differences in inter-country adoption patterns for different technologies. Although Comin et al. separate between the intensive and extensive margins they do not analytically distinguish between these nor analyse the determinants of international diffusion including the possible interactions between the two margins. Comparative studies certainly contribute to our understanding of diffusion patterns but as Stoneman (2002:251) concludes, "diffusion analysis itself has relatively ignored international diffusion especially at the empirical level". It seems to us that the literature lacks testable models in which the link between diffusion paths in different countries is explicitly modelled throughout the diffusion process not just at discrete points (such as the first adoption date, “take-off”, and “saturation”).

1.3.2 Approaches to analysing diffusion

Given the limited literature addressing the issues of main concern, our review indicates that an approach that explicitly takes an international stance is required. The data to enable this is considered in the next section. Here we are concerned with analytical approaches.

The extent of diffusion can be measured in a number of ways, one of which is the proportion of users in a population of potential adopters. Plotting this measure typically generates an S-shaped curve over time as Griliches (1957) observed. The epidemic (and other) approaches to diffusion aim to explain how this pattern is created. The theoretical rationale of epidemic models is that tacit knowledge is required for effective use of a technology and this can only be obtained via personal contact with users. This approach is most common in the marketing literature. In the seminal Bass (1969) model diffusion proceeds because potential adopters learn about the technology from current users (endogenous learning) and some central information source e.g. advertising (exogenous learning). Mansfield (1961) also argued that learning and information are important however in his model their role is to reduce uncertainty about the profitability of adoption.

Estimating S-shaped diffusion curves enables observed patterns to be compared across countries or technologies and the approach may also be used to predict future changes in diffusion within a country. A common representation of the curve is by the three parameter logistic, the parameters of which have been used to indicate the first adoption date, the speed of diffusion, and the end point or saturation level. A common modelling strategy is to regress estimated diffusion speed on measures of expected profitability, risk and learning in a second step (as in Mansfield 1968). The saturation level can also be estimated but typically a particular value is assumed (e.g. Perkins and Neumayer 2005). Analysing the determinants of diffusion in this way can appear rather ad hoc as it does not involve theoretical modelling of the decision to adopt.[5]

In contrast, decision-theoretic approaches take the individual adopter's adoption decision as the starting point. A model is described as decision-theoretic if the adopter (typically a firm) is assumed to choose a profit-maximising level of use at every point in time and usage increases only if this increases (usually as the costs or benefits of adoption change). Much of the literature focuses on the per-period gross benefits, $g_i(t)$, that a potential adopter expects from adoption. The benefits have been modelled as determined by rank, stock, and order effects (Karshenas and Stoneman 1993).

Rank effects are present when potential adopters have different characteristics that affect their ability to realise profit gains and to use the technology effectively. Davies (1979) argued that firm size proxies such differences. The firm's stock of vintage and complementary capital, the availability of other inputs such as skilled labour and

[5] An interesting attempt to give micro-economic foundations for the relationship between information and diffusion is Costa-Font and Mossialos (2005) who study the demand for genetically modified (GM) food. They propose that the less information an individual has about GM food, the more they fear it and the greater their demand for information. A lack of information induces uncertainty related to a new technology which is expected to reduce the speed of diffusion. The study demonstrates that hypotheses typically found in the diffusion literature can be developed using microeconomic reasoning.

own R&D, market power and the growth rate of demand have also been considered (Stoneman 2002:35-6).

If the behaviour of other potential adopters also affects $g_i(t)$ then stock or order effects exist. Both can, but need not, arise from strategic behaviour. In stock models, $g_i(t)$ depends on the extent to which others have adopted the technology at that point in time. The effect is negative if the competitive advantage from using a technology is eroded as competitors also adopt (Karshenas and Stoneman 1993). An example of a game theoretic model with a negative stock effect is Reinganum (1981b) which is the basis of our model in Chapter 4.

The stock effect may in contrast be positive if there are so-called network effects: the benefits of adoption are increasing with the number of users, that is, adoption is more profitable the bigger the network (Saloner and Shepard 1995). The stock effect may therefore be either negative or positive but most commonly the term stock effect is used, as here, for the negative relationship while other terms (such as network effect) are used for the positive relationship.

In order models, $g_i(t)$ depends on the adopter's position in the order of adopters. Ireland and Stoneman (1985) argue that early adopters can obtain prime geographic sites or access to scarce inputs that are complementary to the technology. In Fudenberg and Tirole (1985) order effects arise because of first-mover advantages: early adopters can influence the adoption decisions of others in a way that increases their own profits from adoption. Fudenberg and Tirole argue that if firms cannot commit (to output, adoption) there is an incentive to pre-empt. In their two-player game both players get the same benefit from adoption although they adopt at different points in time. Hoppe (2002) surveys later models in which there is pre-emption but payoffs are not equalised. One such class of models concerns quality-improving technological progress where the product produced with the technology has a higher quality if the firm adopts the technology later (see Dutta et al. 1995).

Empirical evidence (mainly from national or industry-level studies) for rank effects is plentiful. Epidemic effects are also typically found, but stock and order effects are usually difficult to identify in empirical models. Karshenas and Stoneman (1993) presented an encompassing model with epidemic, rank, stock, and order effects which has been used in many empirical studies. A considerable challenge is to find different measures for stock, order and epidemic factors; the current extent of use at the time of adoption is typically used for all three. This is problematic because the relationship with further use is negative according to the stock and order hypotheses and positive for epidemic effects. Inference is impossible in the sense that data can never reject any of the hypotheses, merely suggest that one – usually epidemic effects – is strong enough to "produce" a statistically significant, in this case positive, coefficient. For example, Hollenstein and Woerter (2008) find this positive relationship in e-selling and e-purchasing which, the authors argue, suggests that learning effects are important while stock and order effects "are attenuated by positive network effects that might be quite substantial."[6]

Some studies include other variables which only reflect one of the effects and this increases the potential for inference. In hazard rate models which (typically but not necessarily, see Battisti and Stoneman 2003) study the date of first adoption, the

---

[6] The authors argue that they would be able to disentangle the two effects if time-series data was available.

duration of non-adoption (since the technology became available) can be used to capture epidemic effects as in Hannan and McDowell (1987). However this interpretation of time as an explanatory variable has been criticised by Colombo and Mosconi (1995). They argue that calendar time reflects exogenous factors such as changes in the price and quality of the technology, and that duration should not matter because epidemic effects are captured by current use. As well as calendar time, duration, and an industry-level measure of current use, they include a geographically defined measure of current use to capture epidemic effects. Data from metalworking supports the hypothesis that there is no duration dependence in this set-up. There is evidence of epidemic effects through the geographical measure, but Colombo and Mosconi (1995) conclude that there is no evidence of stock and order effects because the industry measure is insignificant. Another recent and relatively successful attempt to find evidence of stock and order effects is Fusaro's (2009) study of bounce protection schemes adopted by retail banks. Again, the stock of adopters is used to measure both stock and order effects, but Fusaro also includes other measures of the order effect. This allows him to conclude that the results support order effects and are not inconsistent with a stock effect. Negative duration dependence is interpreted as evidence against epidemic effects.

This short review suggests that rank effects have been extensively studied but empirical evidence of stock and order effects is rare because of identifiability problems. In this study, the epidemic approach to diffusion is the focus of Chapter 3 while the stock effect hypothesis is central to Chapters 4 and 5 where rank effects are also considered. In Chapter 5 we also return to the point made by Colombo and Mosconi (1995) that a variable which indicates the passage of time reflects changes in the price of adoption if that is not captured by any other variable in the model.

We do not consider the order effect for which the obvious opportunity arises in Chapter 4. As will be explained below, we develop there a model based on Reinganum (1981a,b). The original model was criticised by Fudenberg and Tirole (1985) who essentially presented an order model as an alternative to a stock model. We have chosen to develop a stock rather than order model because: i) empirical evidence of both is elusive; ii) the stock hypothesis is more straightforward if network effects are not present; iii) in the case of basic oxygen furnaces (Chapter 5), there is no evidence of first-mover advantages in the literature and pre-commitment is not an unreasonable assumption given the size of the investment project.

1.4 Data

Our main data source is an exceptional panel dataset, the Historical Cross-Country Technology Adoption Dataset (HCCTAD), which has been compiled by Diego A. Comin and Bart Hobijn and is freely available at http://www.nber.org/hccta. The availability of data has been one of major obstacles to empirical research (Stoneman 2002) and the HCCTAD provides an excellent opportunity for the study of international diffusion. It features annual country-level data on 21 technologies in 23 developed countries with observations between 1788 and 2001. There is also information on additional country-level variables such as population size, Gross Domestic Product, educational achievement, and other political and social variables. The data is of very good quality and has been collected from widely recognised sources such as Mitchell's (1998) "International Historical Statistics". The panel is unbalanced with some missing observations and we have conducted a limited number of imputations for which details are given in each chapter. In Chapter 2 we investigate the problem of early missing data in some detail and manage it by setting a non-zero threshold which separates users from non-users.

The technologies we examine are: postal services (mail), telephone mainlines, and electricity (all in Chapter 2), the basic oxygen furnace (Chapters 2 and 5) and steam- and motor ships (Chapter 3). Our empirical results are of course specific to HCCTAD but this does not mean that the findings cannot be generalised to other technologies. Our aim has been to keep the empirical analysis at such a level that it is straightforward to replicate the studies with other technologies. Generalisability to a larger number of countries is more questionable because the countries in HCCTAD are all members of the OECD. For each exercise we have initially considered all the countries in HCCTAD but then dropped those for which data coverage is insufficient.[7] The countries that we consider were all members in 1973 (the year New Zealand joined).[8]

In general, inference depends on assumptions about how the sample was drawn and here neither the countries nor the technologies can be considered randomly chosen. We consider the countries to be a sample from the population of OECD countries, and the technologies can all be considered process technologies and it is to these populations that our results apply. The extended version of the database (CHAT) covers 115 technologies (many of which are much more specific than the technologies in HCCTAD) and 150 countries (see Comin et al. 2006). If this was made available, it would be useful to see if our results still hold for this larger sample.

---

[7] E.g. Denmark, Norway and Portugal are excluded from the empirical study in Chapter 5.

[8] The OECD members that we do not consider at all because of missing data are Greece, Iceland, Ireland, Switzerland and Turkey.

The larger sample would most obviously change our measures of international diffusion to include the bigger number of countries. It is likely that this would produce a lower extent of diffusion at each point in time than the current measures because the diffusion of the technologies we have considered here generally happened earlier in the OECD than elsewhere.[9] On the other hand the availability of historical data is generally best for the OECD countries and so missing data would be a bigger problem in the extended sample.

The data was introduced and first used by Comin and Hobijn (2004a) who have since used it and an extended version (known as CHAT) in various analyses (Comin and Hobijn 2004b, 2006; Comin et al. 2006) but not to conduct the type of empirical studies that are reported in this thesis.

[9] For example, the predecessor to the blast oxygen furnace (the open hearth) is still used in some countries although in our sample the switch was completed in 1992.

2 International Diffusion of new technologies: intensive and extensive margins

Equation Chapter 2 Section 1

2.1 Introduction

There is no clear definition in the literature as to precisely what is meant by the international diffusion of new technology, it is however natural to consider that it relates to changes over time in the extent to which world output is produced using, or world consumption is made up of products incorporating, a specific new technology. Examples would include the proportion of cars in the world produced using robots or the proportion of world televisions that incorporate HDTV. We label such measures as indicators of the overall diffusion of new technology. Overall diffusion is the result of two, possibly related, processes. The first concerns the extensive margin and relates to the spread of first use of a new technology across different countries (inter-country diffusion). Thus international diffusion may occur as a new technology is first used by firms or consumers in the United States, then Japan, then France, Germany and so on. The majority of the literature on the international diffusion of new technology is concerned with this extensive margin. The second process concerns the intensive margin and reflects the increasing extent to which the technology is used in different countries post first use (intra-country diffusion)[10]. Most of the literature rarely considers this dimension (e.g. see the review by Keller 2004)[11] however, in terms of welfare it will be the latter process that is most important because it is only as technology is widely disseminated that substantial benefits arise.

[10] Much of the literature considers the study of international diffusion to involve just a comparison of national diffusion paths. That is not, a priori, appropriate.

[11] It is however fair to say also that many national studies of diffusion consider the intensive margin but ignore the extensive.

This distinction between extensive and intensive margins has been picked up by Comin et al. (2006) where, inter alia, they also argue that the intensive margin has been largely ignored. In that paper[12] they analyse a number of related issues such as whether international diffusion is getting faster, whether overall diffusion is logistic and whether there are differences in inter-country adoption patterns for different technologies. Comin et al. (2006) are mostly concerned with the path of overall diffusion whereas our prime concern in this chapter is with the relative importance of inter-country and intra-country diffusion in the overall diffusion process and how this changes as overall diffusion proceeds. The approach taken here is similar to that of Battisti and Stoneman (2003) who explored the relative importance of inter-firm and intra-firm diffusion in overall industry diffusion. That previous exercise illustrated that although inter-firm diffusion was most commonly studied, intra-firm diffusion was in fact the main factor in overall diffusion for most of the study period.

The chapter proceeds by detailing the objectives and methods of analysis in the next section followed by a convenient example with very good data, the international diffusion of postal services, in section 2.3. We then extend to other technologies in section 2.4, and finally discuss implications and present our conclusions.

2.2 Analytical methods

The prime objective of this analysis is to explore the relative importance of inter-country and intra-country diffusion in the overall diffusion process of a new technology at different stages in that overall process.

---

[12] We may also note that Comin et al. (2006) use a larger data set than that available to us, which is the one used in Comin and Hobijn (2004a,2004b, 2006).

Usage of new technology can be measured in a number of ways, three of which are most common: (i) total usage or ownership in time t, which we label $D^1(t)$; (ii) usage or ownership relative to some total output measure, e.g. Gross Domestic Product, which we label $D^2(t)$; and (iii) total usage relative to some estimated post-diffusion (asymptotic or saturation) level of usage. In this chapter we do not want to become involved in actually estimating diffusion curves and thus concentrate upon the first two measures.

The extent of use in any particular country, intra-country diffusion, is measured by the same indicator as overall diffusion but for the individual country. Defining this as $D^k(i,t)$, k =1,2 for country i, intra-country diffusion may be measured, for example, by usage or ownership relative to GDP in country i.

The appropriate measure of inter-country diffusion is the number (or proportion) of countries that are using the new technology at a level $D^k(i,t)$ in excess of some externally chosen base level D^{k*}. The most obvious choice for D^{k*} would be zero. However, data sources rarely pick up very first usage and there are considerable differences across countries in the level of usage that is first recorded. In order to make our analysis less sensitive to such differences in data availability, it is necessary to choose a positive D^{k*} for each given measure of diffusion (see further discussion below).

Let there be N(t) countries in total, of which M(t) in time t are users in the sense that intra-country diffusion exceeds D^{k*}. Define x(i,t) as total usage or ownership in country i at time t. If overall diffusion is to be measured by total usage or ownership then diffusion is simply the sum of x(i,t) across all M(t) using countries. Defining this sum as X(t), overall diffusion is given by

$$D^1(t) = \sum_{i=1}^{M(t)} x(i,t) = X(t) \tag{2.1}$$

which can be written as

$$D^1(t) = M(t) * \frac{X(t)}{M(t)} \tag{2.2}$$

with M(t) being an absolute measure of inter-country diffusion and X(t)/M(t) a measure of average intra-country diffusion equal to the average level of use across the using countries.

Alternatively, if overall diffusion is measured by usage or ownership relative to some measure such as GDP or total output, overall diffusion $D^2(t)$ will be given by total usage across all (using) countries X(t) relative to total output produced in all N(t) countries

$$D^2(t) = \frac{\sum_{i=1}^{M(t)} x(i,t)}{\sum_{i=1}^{N(t)} y(i,t)} \tag{2.3}$$

where y(i,t) is output of country i at time t. Denoting total output of countries in the sample by Y(t) overall diffusion is given by

$$D^2(t) = \frac{X(t)}{Y(t)} \tag{2.4}$$

which may be written as

$$D^2(t) = \frac{M(t)}{N(t)} * \frac{X(t)/M(t)}{Y(t)/N(t)} \tag{2.5}$$

Here, M(t)/N(t) is a measure of the proportion of countries using the technology, an obvious inter-country measure and [X(t)/M(t)] / [Y(t)/N(t)] is average usage in using countries relative to the average output of all countries, a not quite so obvious intra-country diffusion measure. Thus for both $D^1(t)$ and $D^2(t)$ overall diffusion reflects two multiplicative indicators reflecting (i) the number or proportion of using countries and (ii) the average intensity of use in each country.

Denoting measures of inter-country diffusion by z(t) and measures of intra-country diffusion by w(t) we may express (2.2) and (2.5) as relationships between growth rates rather than levels by taking natural logarithms and differentiating with respect to time:

$$\frac{d \ln D^k(t)}{dt} = \frac{d \ln z(t)}{dt} + \frac{d \ln w(t)}{dt} \tag{2.6}$$

The discrete time analogue is

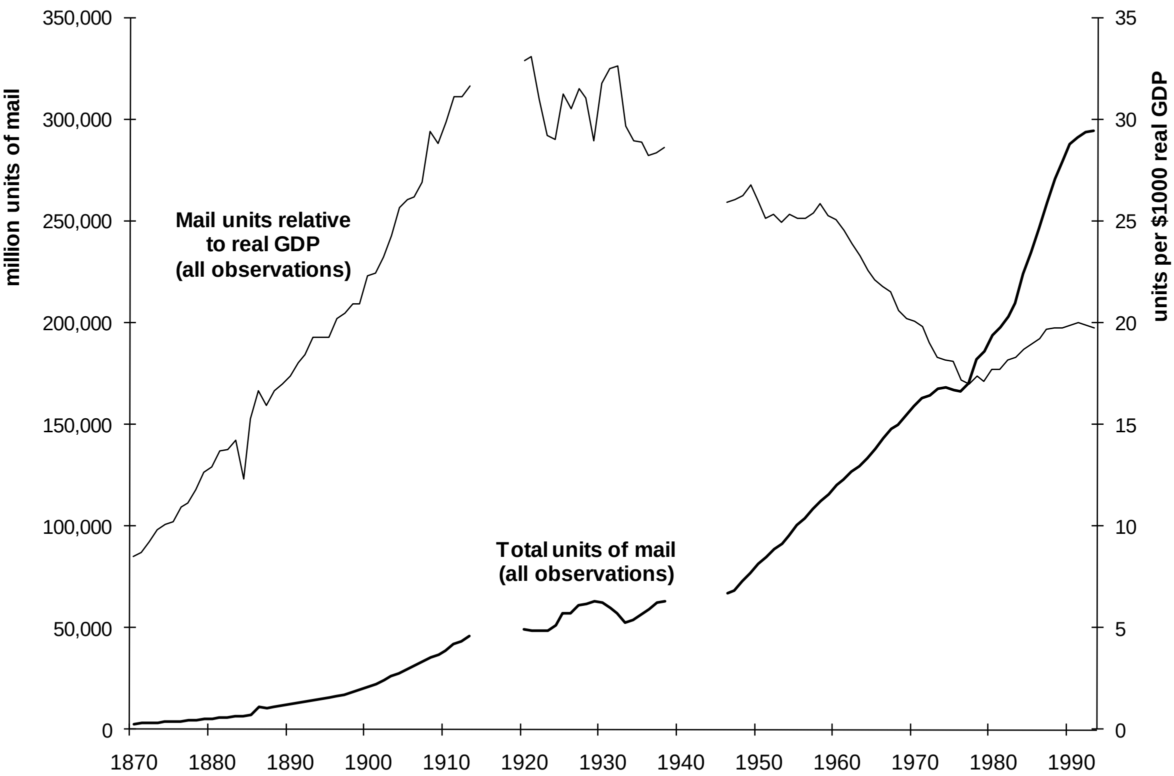
$$\Delta \ln D^k(t) = \Delta \ln z(t) + \Delta \ln w(t) \tag{2.7}$$

Looking at growth in overall diffusion over a time period in this manner allows the analysis of the relative contributions of inter- and intra-country diffusion (Δ ln z(t) and Δ ln w(t) respectively) to overall diffusion.

2.3 A first example

We take postal services as our first technology of interest primarily because the data is good and extensive. The HCCTAD provides annual data on the units of mail handled in up to 21 countries over the period 1830-1993. There are a considerable number of early observations (the earliest are for France and Austria in 1830) and from 1860 onwards annual figures are available for 15 countries or more without many consecutive missing observations. We consider two measures of diffusion, first the units of mail handled and second mail relative to GDP. Annual figures for GDP are available since 1870, but we also have information on GDP for the year 1850 for several countries. Until 1870 the amount of mail handled overall was generally low, and most of the growth in both the level of mail and the level relative to GDP has occurred since then. In Figure 2.1 we plot estimates of these two measures of diffusion since 1870. For this exercise we set D* equal to zero and the impact of missing observations has been smoothed out. This differs from our procedures below but serves to illustrate some of the differences between the two measures of diffusion. There has been a considerable increase in both total mail and mail relative to GDP, but the latter seems to have passed its peak by 1993 while the former continued to increase.

Figure 2.1 Total sample usage of postal services 1870-1993



To undertake a formal separation between the importance of inter- and intra-country effects in the illustrated overall diffusion process, as per the last section, we need first to consider some issues of data availability. The first issue is that for many of the countries in HCCTAD the first observation does not correspond with the very beginning of the diffusion process. This is a frequent occurrence in diffusion studies but is not necessarily a problem if it can be assumed that the level of usage in the unobserved period is below the arbitrary threshold level D*; that is, such countries in this period were effectively non-users. To this end we experiment with different values of D* for each measure of diffusion. An appropriate choice has to be low enough to justify the interpretation of D* as distinguishing users of a technology from non-users, but not so low that we are left with data on users alone (implying that no changes in inter-country diffusion can be captured). Also, an appropriate value should be high enough that the given measure of inter-country diffusion (equal to or a function of the number of users M(t)), $D^k(i,t)$, would be initially relatively low, and low enough that even countries in which usage never reaches a very high level can be considered users of the technology.

The second issue is that there are some countries (namely the United States, Japan, and Ireland) for which the level of usage at the first observation is so high that the countries cannot plausibly be regarded as non-users prior to that date but must be excluded from the analysis until that first observation.[13] Therefore we face the task of deciding which of the 21 countries in the HCCTAD can be included in the sample (determining N(t)). Clearly, with missing observations the sample size cannot be 21 for the whole period of analysis. The alternative of fixing N(t) at the number of countries for which data is available in say 1850 is also not attractive because the sample would be insufficient to represent "international" usage and more importantly because we would not capture the inter-country spread of technology over time. Finally, to let N(t) vary with the availability of data would imply that changes in (especially inter-country) diffusion would reflect increases in the availability of data over time. Fortunately, because our concern is with *changes* in overall, inter- and intra-country diffusion over a given time period, we can allow the sample size to vary between periods as long as it is kept fixed within each period. Such a measure of N(t) is also valid because we are interested in the relative (rather

[13] In particular, the earliest data for the United States is from 1886 (at 3747 million units) and Japan does not appear in the data set until 1902 (911 million), while e.g. for Greece we have a first figure of 800 000 units in 1840.

than absolute) contributions of inter- and intra-country diffusion to changes in overall diffusion.

Measuring the diffusion of postal services by total usage or ownership, i.e. $D^1(t)$ as defined above, we have decomposed the growth in overall diffusion for the period 1850-1990 using two alternative values for D^{1*} of 10 million and 50 million units of mail handled. With D^{1*} equal to 10 million we have sufficient data for 15 countries at the start of the period (i.e. N(t)=15). In 8 of these countries the amount of mail handled exceeded 10 million in 1850.[14] Initial overall diffusion (i.e. the total amount of mail handled in these 8 countries) was 812 million units and initial intra-country diffusion (i.e. average usage) was 102 million units. In 1990 all 15 countries were users with the overall level of diffusion equal to 82,954 million units.

Applying equation (2.7) we obtain that 13.6% of the growth in overall diffusion was due to an increase in inter-country diffusion (that is an increase in the number of users M(t) from 8 to 15) and 86.4% was due to higher intra-country diffusion (that is an increase in average usage from 102 million to 5,530 million units). Taking D^{1*} equal to 50 million units the growth in overall diffusion between 1850 and 1990 decomposes such that 35.9% can be attributed to inter-country diffusion and 64.1% to intra-country diffusion.[15] This example suggests, not surprisingly, that in the long run, overall diffusion is primarily driven by an increasing intensity of usage within using countries.

---

[14] The countries in the sample are: Austria, Belgium, France, Germany, Netherlands, Spain, Switzerland, United Kingdom (the users in 1850) and Australia, Denmark, Finland, Greece, New Zealand, Norway, Sweden (who become users during the period).

[15] With D^{1*}=50 million the sample consists of 17 countries 3 of which were users in 1850.

Omitting the war years 1938-1950 because the amount of mail appears very volatile and several countries do not report any figures at all during these years[16] we have conducted the decomposition exercise as described above for each of the decades 1830-1990. The data is presented in Table 2.1, the two panels corresponding to D^{1*} equal to 10 and 50 million units respectively. As described above, we allow N(t) to vary across time but keep it fixed within each decade. N(t) is higher for the higher D^{1*} because we include some countries with no data as 'non-users' – these can be assumed to have a level of usage below 50 million (but not below 10 million). Missing data is approximated by values for a year in close proximity to the start of each decade where available, or by assuming linear growth if there are several consecutive missing observations.

[16] The extreme case is Canada for which no data is available for 1915-1947.

Table 2.1 Changes in overall diffusion $D^1(t)$, inter-country diffusion z(t) and intra-country diffusion w(t)

Panel A. D*=10 million units handled

Time period	Sample size	Level of diffusion						Growth	Share of growth	
		Overall (million units)		Inter-country (users)		Intra-country (million units)		of overall diffusion	inter-country	intra-country
t, t+10	N	D(t)	D(t+10)	z(t)	z(t+10)	w(t)	w(t+10)	Δlog D(t)	Δlog z(t) /Δlog D(t)	Δlog w(t) /Δlog D(t)
1830-1840	10	127	182	2	2	64	91	0.357	0%	100%
1840-1850	11	353	668	3	4	118	167	0.639	45.1%	54.9%
1850-1860	15	812	1,585	8	9	102	176	0.669	17.6%	82.4%
1860-1870	16	1,693	2,892	10	12	169	241	0.535	34.1%	65.9%
1870-1880	17	2,917	5,311	13	15	224	354	0.599	23.9%	76.1%
1880-1890	18	5,339	8,563	16	18	334	476	0.472	24.9%	75.1%
1890-1900	19	12,568	21,056	19	19	661	1,108	0.516	0%	100%
1900-1910	20	21,967	38,493	20	20	1,098	1,925	0.561	0%	100%
1910-1920	20	38,493	49,717	20	20	1,925	2,486	0.256	0%	100%
1920-1930	20	49,172	62,515	20	20	2,459	3,126	0.240	0%	100%
1930-1938	20	62,515	64,012	20	20	3,126	3,201	0.024	0%	100%
1950-1960	21	81,268	120,088	21	21	3,870	5,718	0.390	0%	100%
1960-1970	21	120,088	158,748	21	21	5,718	7,559	0.279	0%	100%
1970-1980	21	158,748	194,042	21	21	7,559	9,240	0.201	0%	100%
1980-1990	20	187,875	281,637	20	20	9,394	14,082	0.405	0%	100%

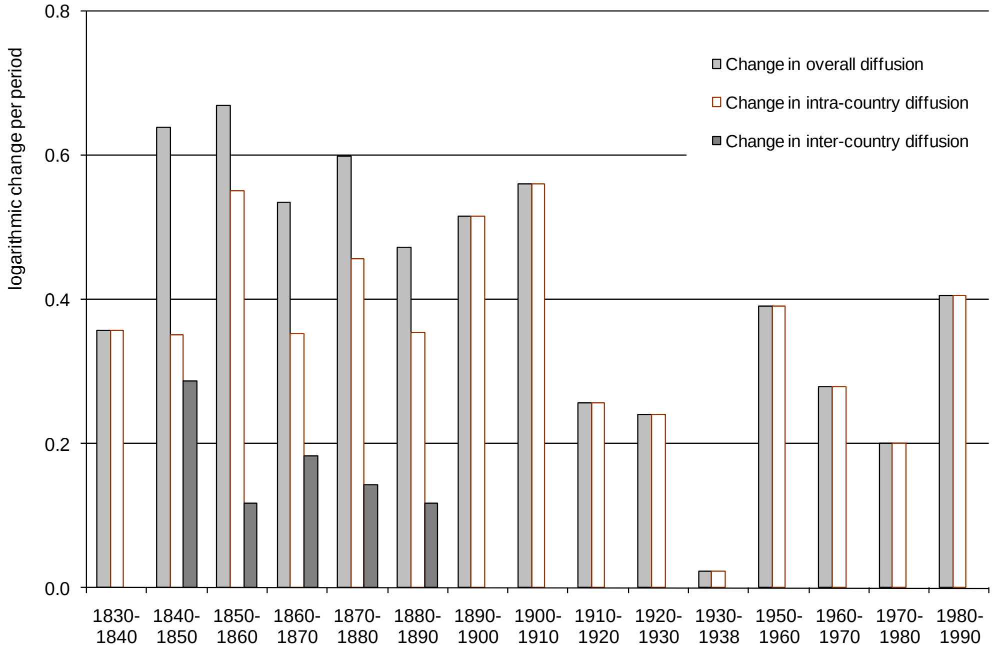
Panel B. D*=50 million units handled

Time period	Sample size	Level of diffusion						Growth	Share of growth	
		Overall (million units)		Inter-country (users)		Intra-country (million units)		of overall diffusion	inter-country	intra-country
t, t+10	N	D(t)	D(t+10)	z(t)	z(t+10)	w(t)	w(t+10)	Δlog D(t)	Δlog z(t) /Δlog D(t)	Δlog w(t) /Δlog D(t)
1830-1840	14	115	163	1	1	115	163	0.346	0%	100%
1840-1850	16	334	632	2	2	167	316	0.637	0%	100%
1850-1860	17	718	1,500	3	6	239	250	0.737	94.0%	6.0%
1860-1870	18	1,608	2,836	7	9	230	315	0.567	44.3%	55.7%
1870-1880	18	2,836	5,211	9	11	315	474	0.608	33.0%	67.0%
1880-1890	18	5,211	8,417	11	13	474	647	0.480	34.8%	65.2%
1890-1900	19	12,422	21,017	14	17	887	1,236	0.526	36.9%	63.1%
1900-1910	20	21,928	38,462	18	19	1,218	2,024	0.562	9.6%	90.4%
1910-1920	20	38,462	49,717	19	20	2,024	2,486	0.257	20.0%	80.0%
1920-1930	20	49,172	62,515	20	20	2,459	3,126	0.240	0%	100%
1930-1938	20	62,515	64,012	20	20	3,126	3,201	0.024	0%	100%
1950-1960	21	81,268	120,088	21	21	3,870	5,718	0.390	0%	100%
1960-1970	21	120,088	158,748	21	21	5,718	7,559	0.279	0%	100%
1970-1980	21	158,748	194,042	21	21	7,559	9,240	0.201	0%	100%
1980-1990	20	187,875	281,637	20	20	9,394	14,082	0.405	0%	100%

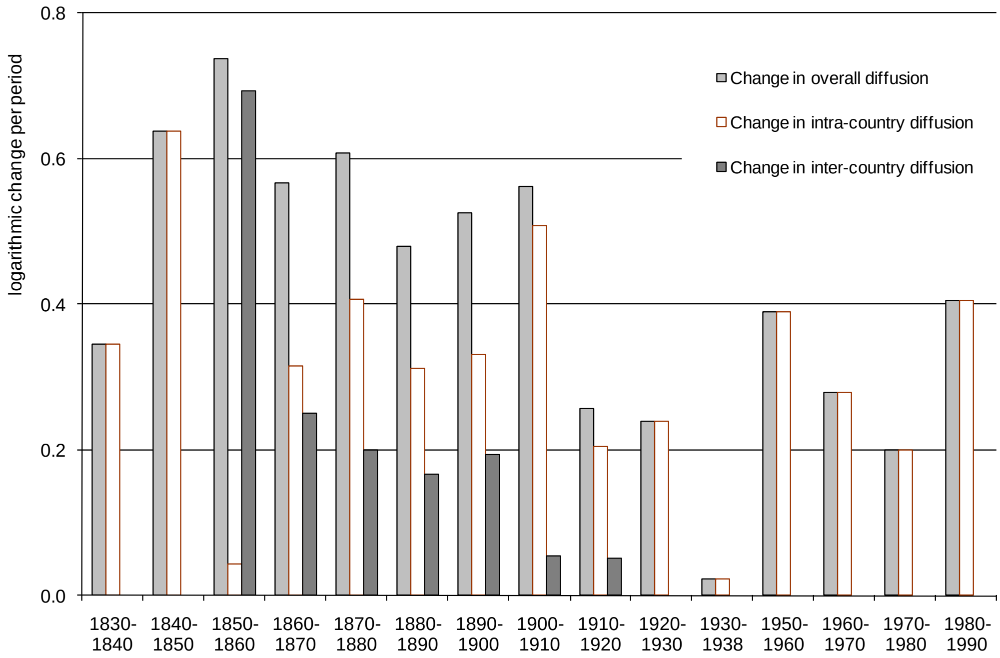
Notes: Overall diffusion $D^1(t)$ is measured by million units of mail handled in all countries M(t) that are users, X(t). Inter-country diffusion z(t) is the number M(t) computed for D^{1*} equal to 10 million and 50 million units of mail in panels A and B respectively. Intra-country diffusion w(t) is computed as the average amount of mail in all using countries measured as X(t)/M(t).

Figure 2.2 Decomposition of changes in $D^1(t)$ by decade: postal services

Panel A. D*=10 million units handled



Panel B. D*=50 million units handled



Notes: Change in overall diffusion is computed from the data in Table 2.1 as log $D^1(t+10)$ – log $D^1(t)$. Change in inter-country diffusion is log $M(t+10)$ – log $M(t)$. Change in intra-country diffusion is log $[X(t+10)/M(t+10)]$ – log $[X(t)/M(t)]$.

Table 2.1 reveals that overall diffusion and average usage have increased continuously throughout the period. The growth in the decade 1980-1990 is particularly striking because our data does not include telefaxes or electronic mail, which, as substitute technologies, we anticipate would reduce demand for traditional mail. There were no increases in inter-country diffusion after 1890 (1920) for D^{1*} equal to 10 million (50 million).

Using these estimates and applying equation (2.7) we plot changes in the log of the overall, inter- and intra-country usage measures in Figure 2.2. We observe first that growth in overall diffusion has tended to be smaller in the second half of the observation period. Secondly, we note that inter-country diffusion slowed as the diffusion process proceeded, especially when D^{1*} is set at 50 million units.

In Figure 2.3 we plot the percentage contribution of changes in inter- and intra-country diffusion to the growth of overall diffusion. We observe that after the initial decades in which the number of users was constant, changes in inter-country diffusion accounted for nearly 50 per cent of overall diffusion. As diffusion (and time) proceeded changes in intra-country diffusion began to dominate the overall growth process. The declining importance of inter-country diffusion is especially evident for D^{1*} equal to 50 million after the decade 1850 – 1860 during which 94 per cent of overall growth was due to an increase in the number of users.

Figure 2.3 Relative contributions of changes in z(t) and w(t) on changes in D¹(t)

Panel A. D*=10 million units handled

Panel B. D*=50 million units handled

Notes: The contribution of changes in inter-country diffusion is computed from the data in Table 2.1 as [log z(t+10) – log z(t)] / [log D^1(t+10) – log D^1(t)] for each decade. Similarly the contribution of changes in intra-country diffusion is [log w(t+10) – log w(t)] / [log D^1(t+10) – log D^1(t)].

We now repeat the decomposition exercise using $D^2(t)$ as in equation (2.5) measuring the diffusion of postal services in country i by units of mail handled x(i,t) relative to real GDP y(i,t). From HCCTAD we measure GDP in 1990 international Stone-Geary dollars. A country is defined as a user if x(i,t)/y(i,t) exceeds the base level D^{2*} which we set at 5 units of mail per $1000 real GDP.

In Table 2.2 below we present the data for each of the decades 1850 – 1990 omitting the war years 1938 – 1950 as above. All three measures of diffusion, overall, inter and intra, grew until 1920 (with the exception of intra-country diffusion which declined 1850 – 1860). In the following decades up to but excluding 1980 – 1990, overall diffusion fell because sample average usage was falling, i.e. GDP was growing more than proportionally to the amount of mail. However, in the last decade, 1980 – 1990, the increase in mail was so considerable that our measure of overall diffusion $D^2(t)$ also increased (compare Table 2.1). This occurred despite a reduction in inter-country diffusion (which was due to x(i,t)/y(i,t) dropping below D^{2*} in one country).[17] Changes in logs of overall, inter-, and intra-country diffusion by decade are plotted in Figure 2.4, from which it is immediately clear that overall diffusion grew fastest in the early decades of the study period. Changes in inter-country diffusion follow a similar (and perhaps even more pronounced) pattern to that found above for $D^1(t)$. That is, increases in inter-country diffusion were large in the first three sample decades, but since then the ratio M(t)/N(t) has changed very little if at all.

[17] All changes in intra-country diffusion in 1960-1990 were due to Greece. Greece proved a difficulty for our analysis since usage x(i,t)/y(i,t) was very low throughout the period. We experimented with lower values of D^{2*} but data availability for other countries suggested that 5 units per $1000 was the most appropriate choice.

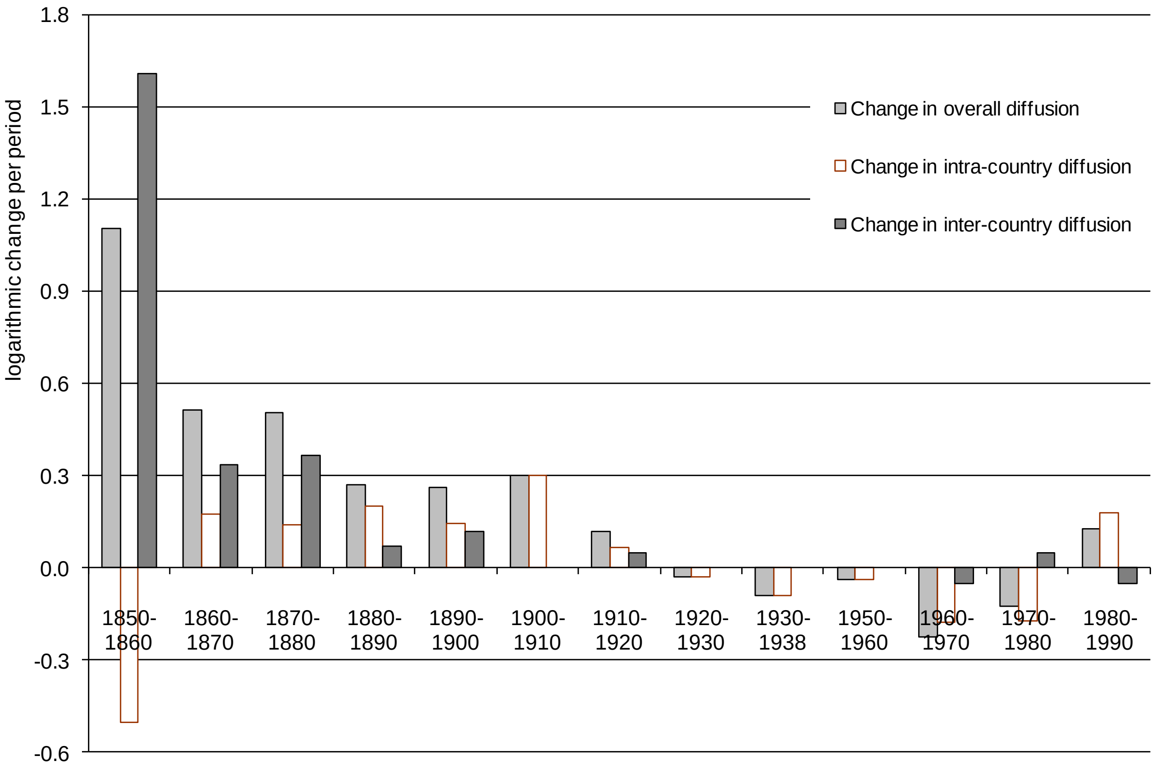
Table 2.2 Changes in overall diffusion $D^2(t)$, inter-country diffusion z(t) and intra-country diffusion w(t)

D*=5 units of mail handled per $1000 real GDP

Time period	Sample size	Level of diffusion						Growth of overall diffusion	Share of growth	
		Overall (mail/$1000 GDP)		Inter-country (user proportion)		Intra-country (mail/$1000 GDP)			inter-country	intra-country
t, t+10	N	D(t)	D(t+10)	z(t)	z(t+10)	w(t)	w(t+10)	Δlog D(t)	Δlog z(t) /Δlog D(t)	Δlog w(t) /Δlog D(t)
1850-1860	13	1.7	5.2	0.08	0.38	22.2	13.4	1.106	145.5%	-45.5%
1860-1870	13	5.2	8.6	0.38	0.54	13.4	16.0	0.513	65.6%	34.4%
1870-1880	17	7.6	12.5	0.53	0.76	14.3	16.4	0.507	72.6%	27.4%
1880-1890	17	12.6	16.5	0.82	0.88	15.3	18.7	0.270	25.5%	74.5%
1890-1900	18	17.2	22.3	0.89	1	19.3	22.3	0.262	44.9%	55.1%
1900-1910	19	22.1	29.9	1	1	22.1	29.9	0.302	0%	100%
1910-1920	20	29.7	33.4	0.95	1	31.2	33.4	0.118	43.6%	56.4%
1920-1930	20	33.6	32.6	1	1	33.6	32.6	-0.031	0%	100%
1930-1938	20	32.6	29.7	1	1	32.6	29.7	-0.090	0%	100%
1950-1960	21	26.1	25.1	1	1	26.1	25.1	-0.040	0%	100%
1960-1970	21	25.1	20.0	1	0.95	25.1	21.0	-0.225	21.7%	78.3%
1970-1980	21	20.0	17.7	0.95	1	21.0	17.7	-0.124	-39.4%	139.4%
1980-1990	20	17.8	20.2	1	0.95	17.8	21.2	0.128	-40.2%	140.2%

Notes: Overall diffusion $D^2(t)$ is measured by units of mail handled per $1000 real GDP in 1990 international Stone-Geary dollars. Inter-country diffusion z(t) is the proportion M(t)/N(t). Intra-country diffusion w(t) is the average amount of mail in all using countries divided by the average real GDP across all countries N(t), that is w(t) = [X(t)/M(t)] / [Y(t)/N(t)] where $X(t) = \sum^{M(t)} x(i,t)$ and $Y(t) = \sum^{N(t)} y(i,t)$.

Figure 2.4 Decomposition of changes in D²(t)



Notes: Change in overall diffusion is computed from the data in Table 2.2 as log $D^2(t+10)$ – log $D^2(t)$. Change in inter-country diffusion is log $(M(t+10)/N)$ – log $(M(t)/N)$ where N is held constant for each decade. Change in intra-country diffusion is $\log\{[X(t+10)/M(t+10)]/[Y(t+10)/N]\} - \log\{[X(t)/M(t)]/[Y(t)/N]\}$.

Finally we plot the percentage contribution of changes in inter- and intra-country diffusion to growth in overall diffusion (Figure 2.5). A very similar pattern emerges as above for $D^1(t)$. Namely, while growth in the number of users was driving changes in overall diffusion in the first three decades (1850-1880) since then it has been the increasing intensity of usage by existing users that has dominated.

Figure 2.5 Relative contributions of changes in z(t) and w(t) on changes in $D^2(t)$

Notes: The contribution of changes in inter-country diffusion is calculated from the data in Table 2 as [log z(t+10) – log z(t)] / [log D^2(t+10) – log D^2(t)] for each decade. Similarly the contribution of changes in intra-country diffusion is [log w(t+10) – log w(t)] / [log D^2(t+10) – log D^2(t)].

2.4 Other technologies

In order to confirm that our findings are not technology specific we have undertaken similar exercises for three other technologies in the HCCTA dataset (although had others been chosen the essence of the results would have been no different). The three are electricity, telephones and the basic oxygen steel making processes. Diffusion in these we measure respectively by megawatt hours of electricity output, the number of telephone lines, and tonnes of steel produced. In each case we have undertaken the analysis using both $D^1(t)$ and $D^2(t)$ diffusion indicators. For brevity however we present only results for $D^2(t)$ measures which look at usage relative to an indicator of total output.

2.4.1 Electricity

In Table 2.3 we present the results relating to the diffusion of electricity measured as megawatt hours of electricity output relative to GDP. D^{2*} is chosen as 0.50 mwhrs per million dollars real GDP. Clearly since 1900 through to the end of the period world electricity output relative to GDP has increased continuously. Inter-country diffusion however was complete in this sample of countries by around 1950 after which all further extensions of use reflected greater intra-country diffusion. This is also illustrated by the fact that, by the 1930s, extensions of intra-country diffusion were already contributing more to overall diffusion than further inter-country diffusion.

Table 2.3 The diffusion of electricity (1900 – 1998)

D^{2*}= 0.50 mwhrs per $1 million real GDP

Time period t, t+10	Sample size N	Electricity output (100 mwhrs) X(t)	X(t+10)	Level of diffusion: Overall (mwhrs/$m GDP) D(t)	D(t+10)	Inter-country (user proportion) z(t)	z(t+10)	Intra-country (mwhrs/$m GDP) w(t)	w(t+10)	Growth of overall diffusion Δlog D(t)	Share of growth: inter-country Δlog z(t) /Δlog D(t)	intra-country Δlog w(t) /Δlog D(t)
1900-1910	14	0	691	0	0.06	0	0.07	-	0.80	-	-	-
1910-1920	15	691	7,744	0.06	0.55	0.07	0.20	0.85	2.76	2.271	48.4%	51.6%
1920-1930	20	9,648	25,803	0.65	1.32	0.40	0.75	1.62	1.75	0.709	88.6%	11.4%
1930-1938	21	25,881	37,216	1.32	1.69	0.76	0.86	1.73	1.97	0.251	46.9%	53.1%
1938-1950	21	37,216	74,083	1.69	2.38	0.86	0.95	1.97	2.50	0.341	30.9%	69.1%
1950-1960	21	74,083	162,223	2.38	3.39	0.95	1	2.50	3.39	0.354	13.8%	86.2%
1960-1970	21	162,223	344,751	3.39	4.35	1	1	3.39	4.35	0.252	0%	100%
1970-1980	21	344,751	529,466	4.35	4.83	1	1	4.35	4.83	0.103	0%	100%
1980-1990	21	529,466	701,783	4.83	4.85	1	1	4.83	4.85	0.006	0%	100%
1990-1998	21	701,783	827,901	4.85	5.29	1	1	4.85	5.29	0.085	0%	100%

2.4.2 Fixed line telephony

In Table 2.4 we present the data relating to fixed line telephony measuring diffusion by the number of telephone lines relative to GDP. D^{2*} is chosen as 1 mainland telephone per one million dollars real GDP. Once again overall diffusion, beginning around 1900, has been continuously extending. However the inter-country spread was complete by 1930, and all growth beyond that date has been through extending intra-country usage.

Table 2.4 The diffusion of mainland telephones (1890 – 1998)

D^{2*}= 1 mainland telephone line per $1 million real GDP

				Level of diffusion						Growth	Share of growth	
Time period	Sample size	Telephones (1000 lines)		Overall (lines/$m GDP)		Inter-country (user proportion)		Intra-country (lines/$m GDP)		of overall diffusion	inter-country Δlog z(t)	intra-country Δlog w(t)
t, t+10	N	X(t)	X(t+10)	D(t)	D(t+10)	z(t)	z(t+10)	w(t)	w(t+10)	Δlog D(t)	/Δlog D(t)	/Δlog D(t)
1890-1900	13	0	1,062	0	1.1	0	0.46	0	2.4	-	-	-
1900-1910	16	1,133	5,555	1.2	4.5	0.56	0.75	2.1	5.9	1.336	21.5%	78.5%
1910-1920	18	5,746	11,288	4.5	7.7	0.72	0.83	6.2	9.2	0.536	26.7%	73.3%
1920-1930	21	11,329	19,301	7.6	9.8	0.86	1	8.8	10.3	0.259	40.6%	59.4%
1930-1938	21	19,301	22,167	9.8	10.1	1	1	10.3	10.6	0.026	0%	100%
1938-1950	21	22,200	39,576	10.1	12.7	1	1	10.1	12.7	0.231	0%	100%
1950-1960	21	39,576	73,290	12.7	15.3	1	1	12.7	15.3	0.186	0%	100%
1960-1970	21	73,290	139,782	15.3	17.7	1	1	15.3	17.7	0.143	0%	100%
1970-1980	21	139,782	256,262	17.7	23.4	1	1	17.7	23.4	0.280	0%	100%
1980-1990	21	256,262	376,995	23.4	26.1	1	1	23.4	26.1	0.110	0%	100%
1990-1998	21	376,995	485,188	26.1	28.2	1	1	26.1	28.2	0.079	0%	100%

2.4.3 The basic oxygen steel-making process

Finally we look at usage of the basic oxygen steel making process, this time conducting the exercise over 5-year periods. Here we measure diffusion by the proportion of all crude steel produced using the basic oxygen furnace with D^{2*} chosen as 10 per cent. The data is reproduced in Table 2.5. The revealed pattern is now the familiar one. By 1970 inter-country diffusion had been completed and all diffusion after that date reflected increased intra-country usage. In this case however overall diffusion had peaked in the late 1980s after which the extent of usage declined. That decline was the result of declining intra-country usage in 1985 – 1995 and the abandonment of the technology by one country (Luxembourg) after 1997.

Table 2.5 The diffusion of the basic oxygen process (1960 - 2000)

D^{2*}= 10% of crude steel produced using the basic oxygen furnace

Time period t, t+5	Sample size N	Basic oxygen output (1000 tonnes) X(t)	X(t+5)	Level of diffusion: Overall (% oxygen) D(t)	D(t+5)	Inter-country (user proportion) z(t)	z(t+5)	Intra-country (% oxygen) w(t)	w(t+5)	Growth of overall diffusion Δlog D(t)	Share of growth: inter-country Δlog z(t) /Δlog D(t)	intra-country Δlog w(t) /Δlog D(t)
1960-1965	13	5,037	68,843	2%	24%	0.23	0.92	10%	26%	2.347	59.1%	40.9%
1965-1970	13	68,843	198,851	24%	54%	0.92	1	26%	54%	0.806	9.9%	90.1%
1970-1975	15	205,628	250,975	53%	67%	1	1	53%	67%	0.235	0%	100%
1975-1980	16	251,276	259,688	67%	69%	1	1	67%	69%	0.017	0%	100%
1980-1985	16	259,688	243,208	69%	69%	1	1	69%	69%	0.010	0%	100%
1985-1990	16	243,208	246,020	69%	67%	1	1	69%	67%	-0.030	0%	100%
1990-1995	16	246,020	232,303	67%	62%	1	1	67%	62%	-0.079	0%	100%
1995-2000	16	232,303	243,836	62%	62%	1	0.94	62%	66%	-0.001	6699%	-6599%

2.5 Conclusions, limitations and implications

We have argued that international diffusion involves two margins – the extensive and the intensive. The former reflects usage extending to previously non-using countries, the latter refers to increasing usage in countries post first use. The majority of the literature on international diffusion considers only the extensive margin. We have shown however that the extensive margin only plays a major role in international diffusion in the early years of the diffusion process. In the later years it is the intensive margin that is important. Thus in the early part of the diffusion process the inter-country spread of a technology is the more important in the diffusion process while in the later years it is mainly intra-country diffusion that is important. This is equivalent to the findings of Battisti and Stoneman (2003) that in the early stages inter-firm diffusion is most important in industry diffusion but in later stages intra-firm diffusion dominates.

The relative importance of the internal and external margins suggests how one may better compare relative diffusion performance across different countries (for an earlier approach to this see Canepa and Stoneman, 2004). The above is only an accounting exercise and so provides no information upon the forces that drive diffusion. We consider of prime importance to be how the internal and external margins are linked. In particular: (i) is intra-country diffusion affected by inter-country diffusion or the extensive margin? – a question never asked in the extensive body of domestic diffusion studies as far as we are aware; and (ii) is inter-country diffusion affected by intra-country diffusion or the intensive margin? – again a question we do not recall seeing before. These seem to us to be crucial questions in understanding the overall diffusion process.

In the next three chapters we look at how the two main theoretical approaches to diffusion, epidemic and decision-theoretic models, can be used to establish a relationship between the intensive and extensive margins. The argument we want to put forward is that domestic diffusion, that is, the intra-country process within a given country, cannot be analysed in isolation of the diffusion processes that are taking place elsewhere. We call this the international effect: a relationship between the extent of use elsewhere and diffusion at home. This is a novel approach at looking at international diffusion. The next two chapters make a significant theoretical contribution by making the case for the international effect first within the epidemic then the decision-theoretic modelling frameworks. Specific hypotheses arise which are then tested using empirical examples of the diffusion of a particular technology.

3 An epidemic model of steam- and motor ship diffusion

Equation Chapter 3 Section 1

3.1 Introduction

In this chapter we make a case for a relationship between international and domestic diffusion within the epidemic theoretical framework. We develop a model which nests the classic Bass and Mansfield models where domestic diffusion is only affected by what happens at home; we add to this the effect of international usage. The model is estimated using time-series data on steam- and motor ship diffusion in 15 countries. We find some support for an international diffusion effect, but the fit of the model is not very good. We discuss at length how the time series properties of the model variables may explain these empirical difficulties. However, we also argue that the lack of robustness here serves as a motivation to examine decision-theoretic models of diffusion, which is the subject of the next two chapters.

In the epidemic approach to the adoption of new technology, the extent of use increases as information about the existence of the technology and of its characteristics becomes more widespread. We focus here on two seminal pieces of work in the literature. Bass (1969) argued that potential adopters partly learn about the new technology from those who have already adopted it. When the number of users is small, opportunities to learn are few, but as the number of users increases the flow of information also increases. According to Bass, this spreading of information is the driving force behind extensions in diffusion. His model has been extensively used in the marketing literature to analyse the diffusion of consumer goods. The second seminal work is by Mansfield (1961) who developed a model in which uncertainty about the returns to adoption discourages diffusion. As more information about the technology is obtained, uncertainty about the profitability of the new technology is reduced. This increases the extent of use. Mansfield's work has been influential in the economics literature because of the focus on profitability. In both models, information-spreading is of key importance in diffusion. We refer to Bass and Mansfield's models as the classic models of diffusion.

Our empirical example in this chapter is the diffusion of steam- and motor ships over the period 1809-1939. During this period sailing ships gradually disappeared from world commercial shipping. The diffusion of steamships is our main concern since motor ships did not begin their diffusion until the early 20th century. Unfortunately steam- and motor ship tonnages are not separately reported for most countries in our dataset.[18] Therefore we consider these two together as the new technology that replaces sailing ships.

Steam engines are a General Purpose Technology (Crafts 2004) however steamships have rarely been studied in the diffusion literature. Comin et al. (2006) attempt to classify a large number of technology-country pairs according to the fit of the logistic curve. Using the same data as we here, they fit a logistic curve to two measures: the proportion of steam- and motor ships in total tonnage, and per capita steam- and motor ship tonnage. The authors do not report the logistic parameter estimates for individual technologies but point out that the logistic does not fit the latter measure well because of a "moving ceiling" (Comin et al. 2006:18). Cohn (2005) regresses the yearly change in gross steamship tonnage on the lagged volume of immigration from Europe to the United States. His simple linear regression model is constructed

[18] A separate analysis of steamship diffusion would have been possible only for Belgium, Denmark and Norway. Motor ships are separated from steamships also in figures for Finland, Germany, and the United Kingdom but this separation is incomplete in the early years.

without reference to epidemic or other theoretical models of diffusion. Cohn finds support for a positive relationship between immigration and change in gross tonnage, which he interprets as evidence that shipbuilders responded to demand. He also finds that a dummy for the year 1870 is positive, which he argues reflects the opening of the Suez Canal and the diffusion of the compound engine.

Most studies of steamship diffusion have followed Graham (1956) and Harley (1971) and focused on analysing the gradual technological change in both steam- and sailing ships and changes in the relative prices (typically freight rates) without an empirical model of the diffusion process.

The main technological problem with early steamships was low fuel efficiency. The space taken up by coal in the hull was considerable especially on long journeys, and the price of coal at refuelling stations increased with the distance to Britain which was the main source of coal at the time (Harley 1971). Steamships were initially able to compete with sailing ships only on short journeys and in the transport of passengers and perishable or valuable cargo for which speed and reliability commanded a premium price. The soiling of the bottom of the ship slowed down early steamships while at the same time sailing ships increased their speed with the help of Maury's current and wind charts, published in 1850 (Graham 1956). Soon however speed and reliability became the main advantages of the steamship. Sailing ships were most competitive in the transport of bulk cargo on distant journeys such as from Europe to India, China, Australia and the west coast of America. The Suez Canal, opened in 1869, could not be used by sailing ships which continued to travel around the Cape of Good Hope. Harley (1971) argues that the fuel efficiency of steamships was still so low that this event did not mark a turning point for the 'Age of Sail'.

The main technological improvements in steamship technology were the diffusion of high-pressure engines and developments in metallurgy. The perfection of the high pressure design was reached in stages with fuel efficiency improving at each stage. Steel structures became available from the 1870s which enabled the use of boiler plates and steel tubes that could withstand yet higher pressures. As the price of steel fell to rival that of best quality iron this led to the development of the triple expansion engine in the 1880s (Graham 1956). Over the period of technological improvements the ship size increased and the cost of ships declined. Crew size could be reduced because steel ships were more reliable, required less crew to operate, and were easier to navigate. Altogether these mainly incremental improvements increased the distance margin at which cargo steamships and sailing ships competed on equal terms (Harley 1988).

Benefits of sailing ships included their faster turnaround time and ability to exploit trade winds. Many of the improvements that took place in steamship technology also occurred in sailing ships. In particular, there was an increase in the size and reliability of ships due to the use of iron and steel. Wooden ships were enhanced with metal structures and later iron and steel sailing ships were built. In harbours, the movement of these very large ships was made possible by steam tugs. Capacity per crewman was still higher in sailing ships than steamships in the late 1880s, although the gap was narrowing (Knauerhase 1968). Labour productivity increased due to economies of scale from bigger ships and the higher reliability of steel masts and rigging (Harley 1988). In sailing ships, the use of auxiliary steam engines also increased labour productivity. There is a strand of literature that considers these technological improvements in sailing ship technology to have been an irrational response by sailing ship owners to the threat of the new superior steamship technology. However, Howells (2002) has argued that the historical facts do not support this "sailing ship effect" as it has later become known. He argues that rather than an "irrational" reaction to the threat of steam, the exploitation of iron and steel could as well have been a result of competition among rival sailing ship companies.

We can attempt to distinguish two alternative views regarding why the diffusion of the steamship took as long as it did. The first view is put forward by most studies in economic history, for example Graham (1956), Harley (1971, 1988) and Crafts (2004). According to this view, incremental innovation in steamship technology was crucial for steamship diffusion. As their productivity improved, steamships replaced sailing ships gradually according to the relative cost of shipping by steam and sail. A lower cost was preferred for most cargo. In this view, steamship usage at any given point in time in a particular market is an at an equilibrium or optimal level. The view is supported by the fact that both sailing ship and steamship markets were highly competitive (Broeze 1975, Harley 1971).

The epidemic theoretical framework for the analysis of diffusion suggests that we should examine how information and knowledge about the characteristics of steam- and motor ships spread during the period. This approach regards diffusion as a disequilibrium adjustment process. The fact that some early steamships had sails (Howells 2002) may indicate uncertainty about the benefits of steam; although it may also be interpreted as a response to the unreliability and high costs of early steamships. There is some evidence that early diffusion of steamships in Norway was held back by investors who were uninformed about the true profitability of steamships, and by ship owners who faced capital constraints; however this was not the case in the United Kingdom (Harley 1971, Samstag and Joshi 2005). There appeared to be a degree of suspicion or even prejudice towards steamships by contemporaries; for example, the Lloyds' Register in London viewed steamships with "marked distrust" (Graham 1956:74).

The model we build in this chapter reflects a disequilibrium approach to diffusion, that is, the extent of use at a given point in time is suboptimal except when all sailing ships are replaced by steam- and motor ships. This is so because our objective is to make the case for an international effect within the epidemic theoretical framework, one of the main bodies of diffusion literature. Our objective is not to build a model which explains steam- and motor ship diffusion as such but rather to use this historical case to examine whether there is any support for the international effect hypothesis that arises from the epidemic model. We view our model as a starting point, a framework which can be refined and further developed so that it is more appropriate for the particular historical example. We do not intend to suggest that the technological changes that took place during the period were not important and indeed we argue that market segmentation may explain the lack of robustness in our findings.

Before we proceed it is worth making a note about the deed of nationality of commercial vessels. National shipping, meaning nationally owned and manned commercial vessels, dominated during the 19th and early 20th centuries (Barton 1999). Beginning in the inter-war period but particularly since the Second World War shipping companies have sought to reduce their costs by pursuing shipping registries, so-called "flags of convenience". This has marked the end of national merchant shipping. Examples of convenience flagging for political and military reasons can be found far back in history however it was in the interwar period that the first US and European ships were flagged to Panama for purely economic reasons (Alderton and Winchester 2002). While it would be inappropriate to take countries as the units of analysis for today's shipping industry our study period is before 1939 and so we consider it appropriate to do so.

The structure of the chapter is the following. First, we construct two measures of steam- and motor ship diffusion, one within-country and one world diffusion measure; describe the data and discuss the choice of study periods; and discuss the empirical features of the world diffusion measure. Then, we present the two classic epidemic model of within-country diffusion and our extended model in which the extent of use elsewhere (i.e. international diffusion) is an additional explanatory variable. Third, we estimate these models and discuss the results, including the time series properties of the variables. The data provides some support for the international effect hypothesis however there are several weaknesses that we discuss. We explore the logistic transformation of the diffusion measure as a possible solution to the problem of nonstationarity in the model variables. However, this does not seem to solve the problem in our data. Section A.4 in the Appendix contains a detailed account of the exercise and results of unit root tests on the transformed variables.

3.2 International diffusion of steam- and motor ships

3.2.1 Measuring diffusion

Diffusion studies use a variety of different measures depending on the nature of the particular technology, such as product or process, as well as the objectives of the researcher and the data available to them. In economic history, tonnage figures are most commonly used to analyse the shipping industry. We have data on net registered capacity, which is an estimate obtained by subtracting from the total enclosed space (i.e. gross capacity) the portion devoted to engines, crew's quarters etc.[19] The number of registered ships is also available but problematic because the average steamship capacity changed considerably during the period. We regard steam- and motor ships as a process rather than product technology and therefore a measure of the cargo-carrying capacity is justified.

We face a second choice namely whether to use a level or a proportional measure of the extent of use. This choice was discussed in section 2.2 in the context of international diffusion, and we labelled the alternative measures $D^1(t)$ and $D^2(t)$. In the shipping context, an obvious level measure is the tonnage of steam- and motor ships in a given country and a proportional measure is for example the share of steam- and motor ships in the country's total tonnage. This share is also a measure of the displacement of sailing ships because a country's total stock of merchant ships consists of sailing, steam- and motor ships. We choose this proportional measure because it is precisely this switching process that we are interested in.[20]

---

[19] The net registered capacity is a British method of measurement that was introduced in the mid-1850s and subsequently imitated around the world. One ton is equal to 100 cubic feet or 2.83 cubic meters.

[20] Alternative proportional measures include tonnage per $1000 GDP, for example.

Let T(i,t) denote the total tonnage of commercial ships and let S(i,t) denote the tonnage of commercial steam- and motor ships in country i at time t. The proportion

$$\frac{S(i,t)}{T(i,t)}$$

is our measure of diffusion. The measure does not have units of measurement (for example, tonnage per $1000 real GDP or tonnage per capita) and it is bounded by the values zero and unity. Although S(i,t)/T(i,t) has these particular features this general approach to measuring diffusion could be applied to other technologies especially process technologies.

We then turn to measurement of international diffusion. In chapter 2 we argued that international diffusion can be measured as the simple average of the extent of use in all countries at a point in time. The national diffusion measure S(i,t)/T(i,t) corresponds to x(i,t)/y(i,t) in Chapter 2, that is, it is country i's contribution to a relative measure of international diffusion $D^2(t)$. Let N be the number of countries in the sample; we refer to these countries as the "world". Let W(t) denote the total world steam- and motor ship tonnage at time t and let TW(t) denote the total world tonnage. World diffusion of steam- and motor ships is the ratio of W(t) to TW(t):

$$\frac{W(t)}{TW(t)} = \frac{\sum_{i=1}^{N} S(i,t)}{\sum_{i=1}^{N} T(i,t)}.$$

This corresponds to the relative international diffusion measure $D^2(t)$. Using the terminology of Chapter 2, TW(t) is total production or capital use in the world at time t and W(i,t) is the proportion of that total that embodies steam- or motor ship technology. In our data W(t)/TW(t) tends to unity over time as sailing ships disappear from the world commercial fleet.

3.2.2 Steam- and motor ship data

The data on shipping used in HCCTAD comes from Mitchell (1998). The countries analysed are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Netherlands, New Zealand, Norway, Sweden, United Kingdom, and United States. [21] The study period is 1809-1939. These dates are chosen by us so that the whole process of diffusion is covered in all countries. In practice, data availability presents some constraints to the actual estimation period. We first discuss the start date. In Table 3.1 we have tabulated the first observations of steam- and motor tonnage and of total tonnage for each of the 15 countries. Total tonnage is typically available at an earlier date than steamship tonnage however missing observations are frequently a problem in this early data. Consecutive annual data is available from the year of the first steam- and motor tonnage with the exception of Canada and Finland.[22] Therefore we use this date as the first year of the estimation period in each of the country time-series regressions. The fact that the first steam- and motor ship tonnage is low in all countries except Australia and Canada (see Table 3.1) suggests that the data we have successfully captures even the very beginning of the diffusion process. This is important because estimation of epidemic models is sensitive to any early missing observations.

Although the data is of good quality there was a need for some imputations. We did this by randomly choosing a tonnage level between the values just before and after the missing year. If several consecutive figures were missing, we ordered the random values according to the time trend, that is, in ascending order (over time) for steam-

[21] HCCTAD includes data on Japan but total tonnage and steam and motor tonnage are measured in different ways, so we cannot use this data. More countries are available in Mitchell (1998). However, the quality of this additional data is insufficient for our purposes.
[22] For Canada and Finland the start dates are 1892 and 1873 respectively.

and motor ships and in descending order for sail ships. All years for which more than two consecutive values were imputed are reported in Table 3.1.

Table 3.1 First observations and imputed years

Country	*Steam and motor*		*Total tonnage*		*Imputations*
	first year	*tons*	*first year*	*tons*	*(years)*
Australia	1876	20	1871	97	
Austria	1837	1	1829	188	
Belgium	1837	1	1837	23	1914-1919
Canada	1867*	46	1867	768	
Denmark	1844	1	1829	60	
Finland	1848*	1	1842	120	1888-1891
France	1838	10	1837	697	
Germany	1850	4	1829	265	1914-1922
Italy	1862	10	1862	654	1901-6, 1918-34
Netherlands	1846	2	1846	380	
New Zealand	1870	6	1857	7	
Norway	1866	6	1830	135	
Sweden	1865	12	1830	131	1866-1869
United Kingdom	1815	1	1788	1278	
United States	1809	1	1789	202	

Notes: *Consecutive data for Canada and Finland is available from 1892 and 1873 respectively. For Finland the earlier observations are frequent enough so that we can impute values and include Finland in the international diffusion measure. Imputations in the last column refer to the period following the first observation of steam and motor tonnage for that country. Some further imputations of total tonnage (sailing ships) were made for earlier years to contribute to international diffusion. For Italy, steam- and motor data is missing 1901-6, sailing ship data is missing 1918-23, and both are missing 1926-34.

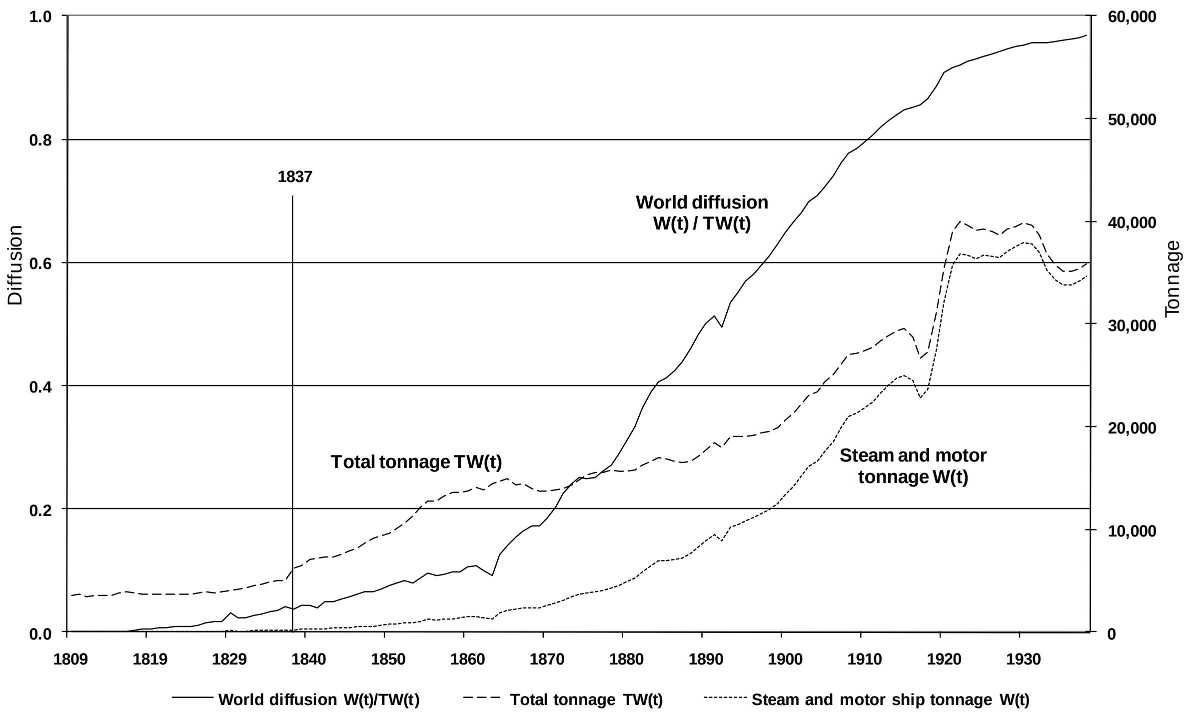
Turning to the world diffusion measure W(t)/TW(t), we chose 1837 as the start date of the estimation period. This choice is determined by the lack of early data on the one hand, and the need to have a measure available as early as possible on the other hand. We have sailing ship tonnage for seven countries in 1837 (Austria, Belgium, Denmark, France, Sweden, United Kingdom, United States) and use data from a nearby year to impute a value for four more countries (Finland, Germany, Netherlands, Norway). Two countries are inferred to have had minimal tonnages (Australia, New Zealand). Imputations are done by taking a random values between two observed years.[23] In cases where there is a strong upwards time trend the imputations are ordered. For Canada and Italy there is too little information even to impute values because the first figures for total tonnage are 768 tons in 1867 and 654 tons in 1862 respectively. For this reason Canada and Italy are excluded from the world diffusion measure.[24]

W(t)/TW(t) is then defined as the average of diffusion in 13 countries (i.e. N=13). The tonnage levels W(t) and TW(t) and world diffusion W(t)/TW(t) are plotted for the period 1809-1938 in Figure 3.1. The measure is also computed from 1809 onwards but missing data is more of a problem in these early years which is reflected in jumps in TW(t) in 1829 and 1837 when more data becomes available.

[23] When the first observation is before 1837, we set the limits for the random values using the two closest observed values. If the first observation is later than 1837, we impute values using the first observation and either a later tonnage if this is lower or some arbitrary lower value which represents approximately a similar growth rate than in the observed data. Note that these imputations are only used for the international diffusion measures.

[24] The other countries can be included for one of the following reasons: 1) all tonnage data is available in 1837 or very close to that date thus it can be imputed; 2) sail tonnage is available and steam- and motor tonnage can be inferred to have been minimal; 3) both are missing but likely to be very small, given the first tonnage that is reported.

Figure 3.1. International diffusion of steam- and motor ships (1809-1938)



The proportional measure of international diffusion has the typical S-shape: the extent of use increases slowly at first then at a higher rate from about the mid-1860s before the rate slows down again as the measure approaches unity. The shape of the diffusion curve suggests that fitting a logistic curve or similar may be appropriate. Most of the increase in diffusion takes place between 1870 and the First World War.[25] In 1837, both W(t)/TW(t) and W(t) are at a low level, at 3.8 per cent or 237 tons. There is an apparent increase in the speed with which usage increases in the mid-1860s, and we may note that this coincides with the diffusion of the high-pressure steam engine. From 1860s until the First World War total world tonnage increases but steam and motor tonnage increases at a faster rate, which produces a

[25] This corresponds to the period in which Crafts (2004) argues steamships were at the peak of their importance for British economic growth.

seemingly steady increase in the proportional measure of diffusion.[26] However, by 1938 there are still some sailing ships in the world commercial fleet.[27]

Some commentators have argued that the First World War hastened the end of the sailing ships (e.g. Samstag and Joshi 2005). We were concerned that there may have been a structural break in the series at this time. A structural break means that the data-generating process is not the same over the whole period, that is, we cannot estimate a single model over the whole period. Although our data concerns merchant ships not military fleet, the strategic role of the industry in the war may have changed the factors that drive diffusion. Perhaps this is evident in how in Denmark and Norway motor ship tonnage increased during the war although sail and steam tonnage fell.[28] We see that there is a noticeable drop in tonnage (both W(t) and TW(t)) during the First World War which reflects the destruction of ships, and that after the war there is a sharp increase in tonnage. In the plot of W(t)/(TW(t) however the war is hardly noticeable, and this finding holds for most of the individual country series i.e. plots of S(i,t)/T(i,t).[29]

We investigate the possibility of a structural break by estimating the model both for the period ending in 1913 and a longer period ending in 1938. The results are used to assess whether there is empirical evidence of a structural break during the First

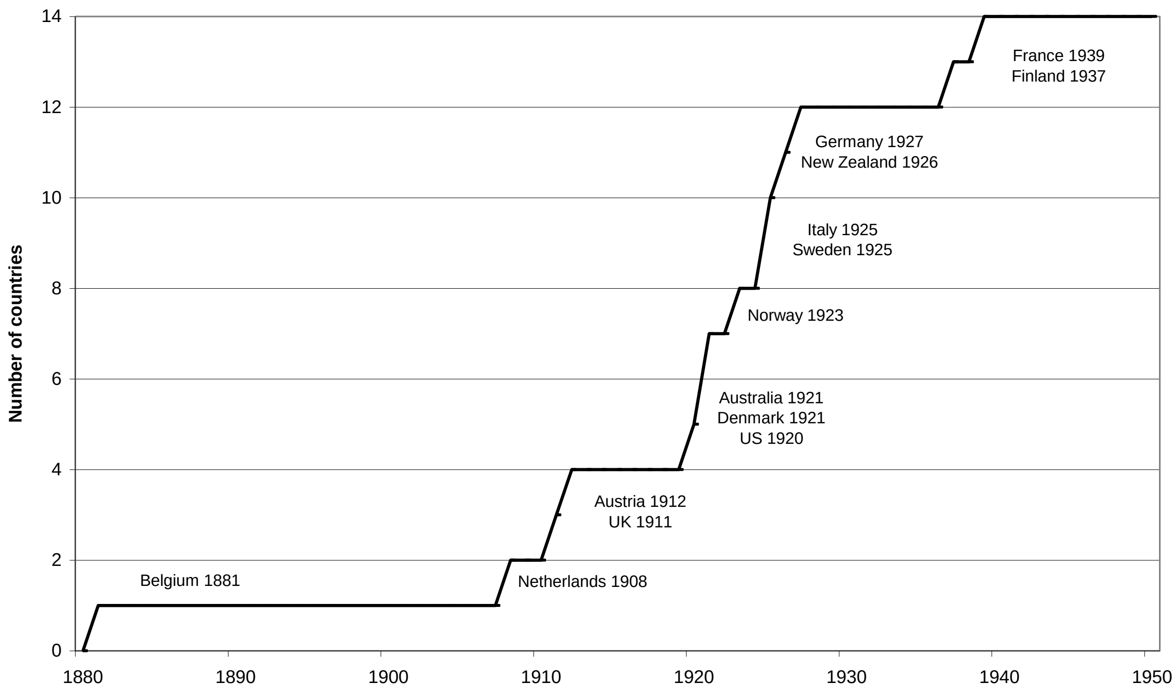
[26] The exception is 1892 which is due to a fall of the order of 1000 tons in the United Kingdom. We suspect that this is a mistake but have been unable to confirm it. The same figure appears in Mitchell (2003).
[27] The sail ship tonnage exceeded 400 tons in Canada and the United Kingdom, 200 tons in the United States and 100 tons in France and Germany. In Britain sailing ships carried small perishable cargoes using small harbours until the 1930s when the use of large ports and lorries for inland transport finally drove sailing ships out of these trades (Greenhill, 1968).
[28] Denmark and Norway are the only countries for which we have data on motor ships during the First World War.
[29] If the contribution of the United Kingdom is taken out of the measure W(t)/TW(t), the war years stand out somewhat more. See Figure 3.3 in Section 3.2.3.

World War. The same concern about a structural break is present regarding the Second World War which is why we do not use data for this period. In Figure 3.2 we indicate for each country the year in which the extent of use reaches 90 per cent; by 1939 this has been reached in all countries except Canada. This suggests that leaving out some of the data post-1939 will not represent a considerable loss of information. Indeed, diffusion had already reached 90 per cent in four countries by 1913 and so for these countries the value of additional information from the inter-war years is likely to be small.

Let us summarise the study period which we use for international diffusion as well as the individual countries. The start date of the study period is the first year in which steam and motor tonnage data is available. This is the value given in the first column of Table 3.1, except that for Canada and Finland the consecutive time series only begin in 1892 and 1873 respectively. For comparison purposes we estimate the models for the United Kingdom and the United States also starting in 1837, which corresponds to the first value in the international diffusion time series. For international diffusion the first year is 1837.

For most countries and for international diffusion we estimate the model over two time periods. The first ends in 1913. The time series is very short for Canada so results are not reported. There is no data available for Austria after 1912, so in these two cases only one model is estimated. The longer estimation period ends in 1939 for most countries. We include data for 1939 because main war activity began in 1940 and indeed the raw data suggests a break in that year. However, there are quite a few countries for which, for various reasons, we do not consider this year to be appropriate. For Italy data is missing for so many consecutive years after 1925 that we prefer not to impute values and end the long period in 1925. (Note that we do impute values for the measure W(t)/TW(t).) We also end the period in 1925 for Australia and in 1931 for Belgium because tonnage falls after these years, and a fall in tonnage is inconsistent with the model assumptions (see below). Finally, we leave out 1939 in the cases of Germany and the Netherlands (no data available) and Norway (tonnage levels jump up considerably) so that the estimation period ends in 1938 for world diffusion, Germany, Netherlands and Norway, and in 1939 for the remaining countries.[30]

Figure 3.2 Year in which steam- and motor ship diffusion reaches 90 per cent (S(i,t)/T(i,t)≥ 0.9)



Notes: Canada is not included because the highest level of diffusion during the study period is 0.66 in 1939.

---

[30] Note that having 1938 as the final observation for world diffusion does not limit the estimation period for individual countries because W(t)/TW(t) enters as a lagged value in the extended model (see below).

3.2.3 International diffusion

We have chosen not to weight the contributions of individual countries to the world diffusion measure. This means that W(t)/TW(t) is dominated by the two countries with the biggest merchant ship tonnage, namely the United Kingdom and the United States. Their combined share was never under 60 percent in either W(t) or TW(t) during the period 1837-1938.[31] The United States initially dominated steamship tonnage contributing over half of the world tonnage until the end of the 1860's. Thereafter the United Kingdom was dominant until the First World War. She had the biggest total as well as steam and motor tonnage in 1870-1919. The share of the United Kingdom in W(t) exceeded 40 per cent in every year during this period except 1919 and 50 per cent in 1874-1903. This was the period in which most of the growth in world steam and motor ship tonnage occurred. A group of the six biggest countries can be identified which constitutes at least 89 per cent of both measures throughout the period 1837-1938. France is initially the third most important country in W(t) until she is overtaken by Germany in 1889, Norway in 1906 and Netherlands in 1921. Figures for the relative importance of the six countries in selected years are given in Table 3.2.

---

[31] The Norwegian steam- and motor ship tonnage increased considerably in 1939 (by 74 percent or over 2000 tonnes) which results in the US and UK shares falling below 60 percent in that year.

Table 3.2 Biggest contributors to W(t) and TW(t) (selected years)

Steam and motor tonnage W(t) and total tonnage TW(t) (tons)

Measure	*1840*	*1860*	*1880*	*1900*	*1920*	*1938*
W(t)	303	1,454	4,846	13,370	32,132	34,697
TW(t)	7,071	13,789	15,688	20,637	35,359	35,848

Contribution of selected countries to W(t)

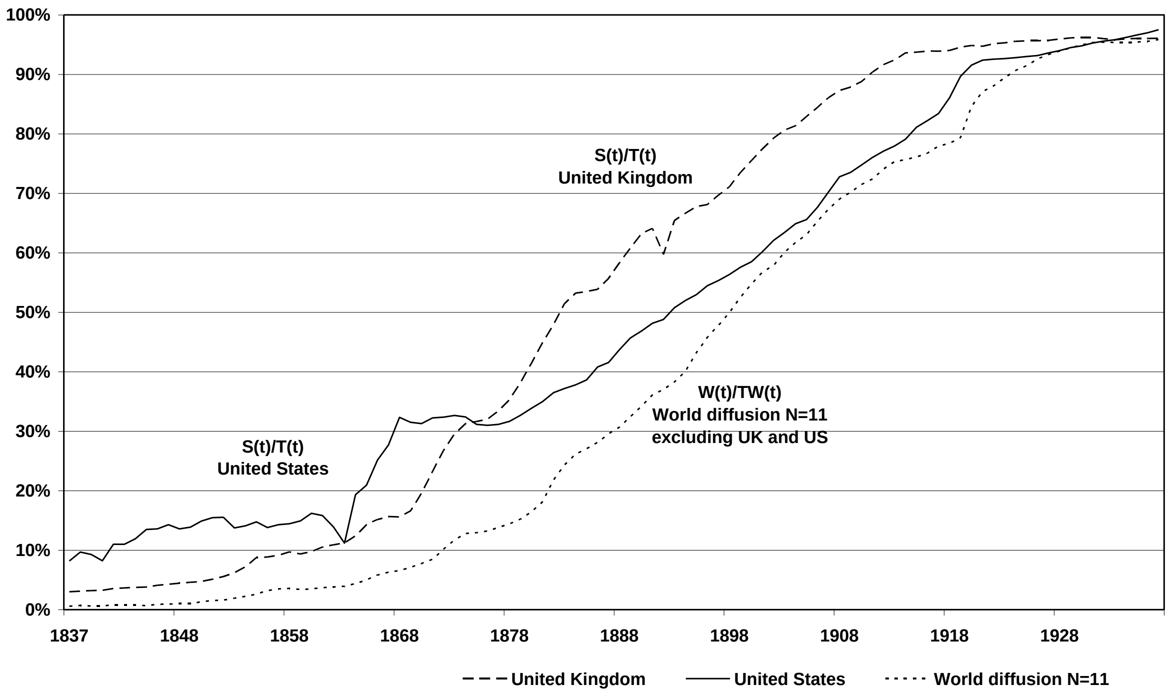
	1840	*1860*	*1880*	*1900*	*1920*	*1938*
France	3.3%	4.7%	5.7%	3.9%	3.4%	4.3%
Germany	0.0%	1.6%	4.0%	9.9%	6.3%	6.8%
Netherlands	0.0%	0.8%	1.3%	2.0%	3.0%	6.4%
Norway	0.0%	0.0%	1.2%	3.8%	4.1%	8.0%
United Kingdom	29.0%	31.2%	56.2%	51.7%	33.5%	29.7%
United States	66.7%	59.7%	25.0%	19.9%	43.0%	34.6%
Total (6 countries)	99.0%	97.9%	93.5%	91.2%	93.4%	89.8%

Contribution of selected countries to TW(t)

Country	*1840*	*1860*	*1880*	*1900*	*1920*	*1938*
France	9.4%	7.2%	5.9%	5.0%	4.3%	4.6%
Germany	5.0%	5.6%	7.5%	9.2%	6.6%	6.9%
Netherlands	4.8%	3.6%	2.1%	1.7%	2.8%	6.2%
Norway	2.9%	3.9%	9.7%	7.3%	4.3%	7.8%
United Kingdom	39.1%	33.8%	41.9%	44.4%	32.1%	29.9%
United States	30.8%	38.8%	22.8%	22.0%	42.7%	34.2%
Total (6 countries)	92.0%	92.9%	89.9%	89.7%	92.8%	89.7%

Domination of W(t)/TW(t) by countries with a large tonnage means that even dramatic changes in diffusion in small countries will not greatly influence world diffusion. Whether or not this property is desirable depends on our theoretical view of how world diffusion affects domestic diffusion. The unweighted measure is consistent with an approach in which the flow of information about the new technology primarily depends on the number of adopters rather than their location. Increases in the extent of use in a country with a small commercial fleet will not give as much information to potential adopters as similar relative increases in a large country because the latter provides a stronger signal. The hypothesis this implies is that international diffusion has a relatively weak effect on domestic diffusion in those countries which contribute most to W(t)/TW(t), that is, in the United Kingdom and the United States. In Figure 3.3, we plot the time path of diffusion in the United Kingdom and United States and also a modified version of international diffusion from which the contribution of these two countries has been subtracted. For most of the study period, domestic diffusion is above the level of use in other countries.

Figure 3.3 Diffusion of steam- and motor ships in the United Kingdom, United States and elsewhere (1837-1938)



An alternative would be to give more weight to some sub-group of countries. In particular, a weighting scheme could be constructed which takes into account the market in which a country's fleet operates. This would require a country-by-country analysis of the relevant factors, the development of a novel weighting scheme since such does not exist in the diffusion literature, and possibly a substantial theoretical argument about the nature of information flows. Models which could be used can be found in the so-called spatial economics literature (discussed in some more detail in section 6.3). Our objective is to conduct a first test of the hypothesis that world diffusion affects domestic diffusion. We consider the unweighted measure to be a good starting point and indeed an important reference point because of its analytical simplicity. The discussion above highlights the need to develop a formal model of the relationship between domestic and international diffusion, which we turn to now.

3.3 Epidemic models of diffusion

3.3.1 Two classic models of diffusion

In this section we discuss how two classical models of intra-country diffusion, that is, models of how the usage of new technology within a country evolves over time, can be applied to the empirical example of steam- and motor ship diffusion. Both models and their variants have been extensively applied in diffusion studies at various levels of aggregation. In general, the country-level has not been as popular as lower levels of aggregation such as the regional, industry or firm-level. For an example of a cross-country study see Gruber and Verboven (2001); for the use of epidemic models in the economics and marketing literatures see Geroski (2000) and for the sociological literature see Strang and Soule (1998).

Epidemic models are disequilibrium models in the sense that changes in diffusion over time are viewed as an adjustment process. The current extent of use is assumed to be below some optimal or target level during the process, and over time usage converges to the optimal level. In models where the optimal level is a constant, say 100 per cent of the population of potential adopters, the optimal level is an asymptotic level in the sense that graphically this is the asymptote to which the diffusion path converges over time. Many models allow the optimal level to be time-dependent and a function of variables such as income of consumers and the price of the technology (see for example Kiiski and Pohjola 2002). We do not make the optimal level a function, but allow it to vary across countries and over time. We denote the optimal level of steam- and motor ship tonnage by S*(i,t).

The Bass (1969) model applied to steam- and motor ship diffusion states that tonnage in country i evolves according to

$$dS(i,t) = \left(\gamma + \beta \frac{S(i,t)}{S^*(i,t)} \right) \left(S^*(i,t) - S(i,t) \right) \tag{3.1}$$

after an initial base of users has been built up. γ and β are positive parameters, and current tonnage S(i,t) is assumed to be smaller than the optimal tonnage S*(i,t). Change in the extent of use dS(i,t) is never negative, that is, it is either constant or increasing in each period. Extensions of use are driven by the discrepancy between current usage and the target level on the one hand (S*(i,t)-S(i,t)) and by information-spreading on the other hand. The proportion who have already adopted, represented by S(i,t)/S*(i,t), are assumed to spread information about the new technology to those who have not yet adopted. The parameter β measures the effect that this information-spreading has on diffusion and it is often called the endogenous speed of diffusion. The second parameter γ is called the exogenous effect because Bass interpreted this as the influence of information flows from (domestic) sources other than current users. Traditionally this role is attributed to advertising.

The Bass model nests another classic model of intra-country diffusion, the Mansfield model. Mansfield (1961) argued that as the number of users increases, uncertainty about the profitability of adoption is reduced which has a positive effect on further diffusion. He also argued that the cost of adoption has a negative effect on diffusion.

These two economic variables are the determinants of diffusion speed, β. The Mansfield model takes the same form as the Bass model except that the parameter γ is equal to zero:

$$dS(i,t) = \beta \frac{S(i,t)}{S^*(i,t)} \left(S^*(i,t) - S(i,t) \right) \tag{3.2}$$

In the Bass and Mansfield models if S*(i,t) is a constant, the time path of S(i,t) follows an S-shape. When current usage is small relative to the asymptotic level, the rate of change dS(i,t) increases with usage up to an inflection point. Above that point, the closer S(i,t) is to S*(i,t) the slower is the rate of change. When usage equals the asymptotic level there is no change in diffusion in that period; if the asymptotic level is a constant S*(i), this occurs when the diffusion process has been completed. Note that neither model can explain why the first adopters adopt in the sense that when S(i,t) is zero dS(i,t) is also zero. Geroski (2000) has pointed out that this is one of the weaknesses of epidemic diffusion models in general.

We need to make an assumption about optimal tonnage S*(i,t), that is, the upper bound of steam- and motor ship diffusion. In the literature, it is not uncommon to assume that S* is some fixed proportion of potential adopters. The proportion can be inferred from values observed in the sample complemented with other information about the technology (e.g. Perkins and Neumayer 2005), or estimated as a coefficient (e.g. Gruber and Verboven 2001). In our case sailing ships disappear completely over time so it is not reasonable to assume that the optimal level is a fixed proportion less than unity. Some studies model S* as a function of other variables. Knudson (1991) for example models the optimal level as a function of the price of output (i.e. of shipping services in our context) and the price of adoption including the price of a complementary input.[32] We make the simple assumption that the population of potential adopters is given by the total tonnage at each point in time:

$$S^*(i,t) = T(i,t) \tag{3.3}$$

This means that the optimal adoption level is 100 per cent of total tonnage throughout the diffusion period. This is the hypothetical “equilibrium”, characterised by all remaining sailing ship owners switching to steam- or motor ships.[33] Total tonnage varies with factors such as the demand for and price of shipping services and our assumption implies that these factors affect S*(i,t) so that the optimal level remains at 100 per cent. Given the improvements in sailing ships as well as in steam- and motor ship technology that are well-established in the literature (see above), the assumption should be considered an approximation not a statement about the true value of S*(i,t). Ideally, we would construct a measure of S*(i,t) for each country from information about the markets (routes) in which the current stock of sailing ships operates, and take into account the improvements in technology over time. However, we do not have sufficient information to do this. We also consider that a more simple approach is appropriate as a starting point and for our purposes.

As argued above the diffusion measure of interest to us is the proportion of steam- and motor ships in total commercial tonnage, S(i,t)/T(i,t). Diving both sides of (3.1)

---

[32] Knudson (1991) treats prices as exogenous to diffusion. Although her empirical results are inconclusive, Knudson’s approach suggests an alternative to the way we proceed here. Namely, we could model S*(i,t) as a function of W(t). This would require a model of the world supply of shipping so that the price of shipping services is endogenous to diffusion, and it would bring the model here closer to the one in Chapter 4. We have not explored this option further because our objective is to maintain the focus on information-spreading. However, modelling the link between international and domestic diffusion in this way is an interesting option for future research.

[33] Note that because the tonnage of ships varies, the assumption does not imply that each sailing ship would be exchanged for exactly one steam- or motor ship.

by T(i,t), substituting S*(i,t)=T(i,t) and applying the standard Euler approach to derive the discrete-time equivalent gives the estimating equation for the Bass model as

$$\frac{S_{i,t} - S_{i,t-1}}{T_{i,t-1}} = \left(\gamma + \beta \frac{S_{i,t-1}}{T_{i,t-1}}\right)\left(1 - \frac{S_{i,t-1}}{T_{i,t-1}}\right) + \varepsilon_{i,t} \tag{3.4}$$

where the errors are assumed to be $\varepsilon_{i,t}$~N(0,σ^2). The dependent variable is the annual change in steam- and motor ship tonnage as a proportion of total tonnage. Changes in diffusion are a function of the proportion of potential adopters – the last term in brackets – multiplied by the information flow from external sources (γ) and the information from current users (β*S(i,t-1)/T(i,t-1)), plus a random term. Estimation of the model requires data on just two variables: the steam- and motor ship tonnage and the sailing ship tonnage in each country. We use non-linear least squares (NLS) as the estimation method.

The estimating equation of the Mansfield model is simply the restricted version of the Bass model with γ=0. If the true value of γ is zero the power of the β estimate from the Mansfield model will be higher (i.e. the standard error smaller).

Although these classic models are rarely applied to international diffusion data, there is no reason why they could not be considered as representations of diffusion in the world as a whole. Using notation from earlier, the Bass model states that world tonnage of steam- and motor ships at time t evolves according to

$$dW(t) = \left(\gamma + \beta \frac{W(t)}{W^*(t)}\right)\left(W^*(t) - W(t)\right) \tag{3.5}$$

Exactly as we did for the intra-country analysis, we first divide both sides by total tonnage TW(t) and then assume that the asymptotic level W*(t) equals the total tonnage TW(t) in all periods. The estimating equation is

$$\frac{W_t - W_{t-1}}{TW_{t-1}} = \left(\gamma + \beta \frac{W_{t-1}}{TW_{t-1}}\right)\left(1 - \frac{W_{t-1}}{TW_{t-1}}\right) + \varepsilon_t \tag{3.6}$$

where $\varepsilon_t \sim N(0,\sigma^2)$. Extensions in world diffusion depend on the current extent of use in the world and the optimal level. Estimating the model requires data on world steam- and motor tonnage and the world sailing ship tonnage. The Mansfield model of world diffusion is the restricted version of the Bass model with $\gamma=0$. The world Bass and Mansfield model results will serve as a benchmark for the analysis of the results from fitting the respective models to data on individual country time-series.

We have presented two classic models of intra-country diffusion, the Bass and the Mansfield models. There is no link between international and domestic diffusion in these because intra-country diffusion is only affected by domestic factors. We now introduce a general model of intra-country diffusion which nests the simple Bass and Mansfield models but extends them to include an international effect.

3.3.2 Extended model with an international effect

In this section we develop an extended model of intra-country diffusion in which there is an international effect: world usage affects domestic diffusion. The model is a general model that nests the Bass and Mansfield models as special cases.

Consider country j which is one of the 13 countries that contribute to the world diffusion measure W(t)/TW(t). We need a measure of diffusion elsewhere, that is, the extent of use in the rest of the world. One option is

$$\frac{W(j,t)}{TW(j,t)} = \frac{\sum_{i=1,i\neq j}^{N} S(i,t)}{\sum_{i=1,i\neq j}^{N} T(i,t)},$$

which is simply the average of diffusion in all N countries except country j. However, we want to write down the model first in terms of the steam- and motor ship tonnage, S(i,t). A relevant measure of use elsewhere is then the difference between the world and domestic steam and motor tonnages, W(t)-S(i,t). For the two countries not included in the world diffusion measure (Canada and Italy), W(t) can be used directly.

Consider how the Bass model could be extended to incorporate world diffusion. A study by Kumar and Krishnan (2002) has attempted something similar. They argue that diffusion in one country is likely to affect another country's diffusion through both of the parameters γ and β. Influence through the "exogenous effect" γ implies that potential adopters get exposed to diffusion as they meet people through travel, read industry publications or newspapers, or see advertisements all of which give information crucial for adoption. In today's information society this argument has a natural appeal but also in the historical shipping context it is reasonable to assume that ship owners had good access to information from around the world (see discussion above). Alternatively, the international effect may be argued to come through the endogenous diffusion speed β so that "the degree to which a potential adopter would place faith on the internally generated information will be affected by what happens in other countries" (Kumar and Krishnan 2002:321). In our context, this second argument may be interpret as that sailing ship owners treat information from domestic steamship owners with suspicion, until world steamship diffusion is at some sufficiently high, 'convincing' level. Both arguments are plausible at least at a hypothetical level. We consider the β effect an interesting alternative if we could also accommodate market segmentation, so that the information value of an adoption depends on the market in which that adopter operates rather than their geographical origin. In our view the "γ channel" is simpler and more intuitive, and it also provides a model that we can estimate using the data available to us.

The hypothesis that arises is that the extent of use elsewhere has a positive effect on domestic diffusion because it increases the amount of information and so reduces uncertainty about steam- and motor ship technology. The general (extended Bass) model is given by

$$dS(i,t) = \left(\gamma + \beta \frac{S(i,t)}{S^*(i,t)} + \alpha \frac{W(t) - S(i,t)}{W^*(t) - S^*(i,t)}\right)\left(S^*(i,t) - S(i,t)\right) \tag{3.7}$$

Here, the extent of use elsewhere enters similarly to the Bass exogenous effect γ. The parameter α is expected to be positive. Restricting α=0 gives the Bass model and α=γ=0 gives the Mansfield model. We are interested in the proportion of steam- and motor ships in total tonnage so we re-write the model as

$$\frac{dS(i,t)}{T(i,t)} = \left(\gamma + \beta \frac{S(i,t)/T(i,t)}{S^*(i,t)/T(i,t)} + \alpha \frac{(W(t) - S(i,t))/(TW(t) - T(i,t))}{(W^*(t) - S^*(i,t))/(TW(t) - T(i,t))}\right)\frac{\left(S^*(i,t) - S(i,t)\right)}{T(i,t)} \tag{3.8}$$

Assuming that S*(i,t)=T(i,t) and W*(t)=TW(t) we have the estimating equation for the general (extended Bass) model

$$\frac{S_{i,t} - S_{i,t-1}}{T_{i,t-1}} = \left(\gamma + \beta \frac{S_{i,t-1}}{T_{i,t-1}} + \alpha \frac{W_{t-1} - S_{i,t-1}}{TW_{t-1} - T_{i,t-1}}\right)\left(1 - \frac{S_{i,t-1}}{T_{i,t-1}}\right) + \varepsilon_{i,t} \tag{3.9}$$

where $\varepsilon_{i,t} \sim N(0,\sigma^2)$.

Nested in this general model are three other models: the Bass and Mansfield models without international effects, and an extended version of the Mansfield model with an international effect. Setting $\gamma=0$ in (3.9) gives the estimating equation for the extended Mansfield model. For the two countries excluded from our world diffusion measure, Canada and Italy, the estimating equations use (W_{t-1}/TW_{t-1}) instead of $(W_{t-1}-S_{i,t-1})/(TW_{t-1}-T_{t-1})$.

The general model (3.9) also nests a model in which diffusion is driven exclusively by world usage: $\beta=0$. There is no corollary to this in the literature but conceptually it is an interesting alternative to the classic epidemic models. In our data, the Bass parameter γ is found to be mostly insignificant and so we estimate a corollary of the Mansfield model; imposing $\gamma=\beta=0$ in the general model we have

$$\frac{S_{i,t}-S_{i,t-1}}{T_{i,t-1}}=\left(\alpha\,\frac{W_{t-1}-S_{i,t-1}}{TW_{t-1}-T_{i,t-1}}\right)\left(1-\frac{S_{i,t-1}}{T_{i,t-1}}\right)+\varepsilon_{i,t} \tag{3.10}$$

where $\varepsilon_{i,t} \sim N(0,\sigma^2)$. We call this the fully coupled model.[34]

3.4 Empirical analysis

3.4.1 Time-series properties

As a first step the time-series properties of the variables are explored; this procedure has become the norm in empirical macroeconomics but not yet in diffusion studies (Stoneman 2002, Battisti and Stoneman 2003). Stationarity is a property that is assumed to hold in conventional distributional results that are applied to the coefficient estimates. If explanatory variables are non-stationary, making inferences based on conventional asymptotic theory for least-squares estimation is inappropriate. Therefore stationarity tests are necessary before estimation of the empirical model.

[34] For Canada and Italy the first term in brackets on the right-hand side is simply $\alpha(W_{t-1}/TW_{t-1})$.

We have presented the general model in its nonlinear form but it is the model's linear form that determines which variables need to be tested for stationarity. The linear version of (3.9) is

$$\frac{S_{i,t}-S_{i,t-1}}{T_{i,t-1}}=\gamma+\left(\beta-\gamma\right)\frac{S_{i,t-1}}{T_{i,t-1}}-\beta\frac{S_{i,t-1}^{2}}{T_{i,t-1}^{2}}+\alpha\frac{W_{t-1}-S_{i,t-1}}{TW_{t-1}-T_{i,t-1}}-\alpha\frac{S_{i,t-1}\left(W_{t-1}-S_{i,t-1}\right)}{T_{i,t-1}\left(TW_{t-1}-T_{i,t-1}\right)}+\varepsilon_{i,t} \quad (3.11)$$

Including the left-hand side variable there are five variables to be tested for stationarity for each of the 15 countries. [35] The tests were conducted for the estimation period, including both a short and a long period where appropriate. Since this model nests all the other models the results are sufficient to make conclusions about the stationarity of variables in all the other models.

For world diffusion we write out equation (3.6), the discrete version of the Bass model of world diffusion, to reveal three more variables that are tested for stationarity:

$$\frac{W_{t}-W_{t-1}}{TW_{t-1}}=\gamma+\left(\beta-\gamma\right)\frac{W_{t-1}}{TW_{t-1}}-\beta\frac{W_{t-1}^{2}}{TW_{t-1}^{2}}+\varepsilon_{i,t} \quad (3.12)$$

The methodology and results of the tests are discussed in detail in the Appendix but we outline the findings here. We are concerned about the possibility that non-rejection is due to structural breaks; in this situation, the Augmented Dickey-Fuller (ADF) test is known to have low power (i.e. likely to not reject a unit root) (Perron 1989). We run a test for a structural break in all those variables for which either the plot of the data or information from some external source suggests that a break might have occurred. If there is evidence of a structural break we then run Perron's

[35] In the cases of Canada and Italy the third and fourth variables on the right-hand side are W(t-1)/TW(t-1) and S(i,t-1)/T(i,t-1) * W(t-1)/TW(t-1) (see above).

(1989) test for a unit root as well as the ADF test. We find that the left-hand side variable is stationary in most cases but the right-hand side variables are not. In particular, the proportion of steam- and motor ships in total tonnage (S(i,t)/T(i,t)) is not stationary even when structural breaks and a linear time trend are taken into account. This means that nonstationarity is a feature of the classic models and so it is also a feature of any general model that nests the classic models.

The reason for non-rejection is likely to be a nonlinear trend in S(i,t)/T(i,t) which is not allowed for in the standard tests for unit roots that we implemented. The Bass model imposes a logistic trend on the data and we investigated whether by assuming a logistic trend we can transform the data so that the problem of nonstationarity is eliminated. If the transformed measure is a stationary variable with a linear time trend, this can provide an alternative model of diffusion in with more desirable time-series properties.

The results of stationarity tests on the logistic transformation of S(i,t)/T(i,t) are discussed in detail in the Appendix. We find that nonstationarity is rejected considerably more often although still not in all time-series. This suggests to us that the nonlinear trend is the most likely explanation for failure to reject nonstationarity. That is, assuming that the time trend is logistic considerably increases the number of countries for which the standard unit root test results are consistent with stationarity. A first step would be to apply some unit root test that allows for a nonlinear time trend; such have been developed in the literature, but these are not as straightforward to implement as the more common ADF and Perron tests. If nonstationarity is still not rejected other transformations such as the Gompertz curve could be considered, and error correction models may provide another route forward.

The time-series properties of the model are not our main concern in this chapter. We recognise that a failure to establish stationarity is a concern however because the logistic transformation is not entirely satisfactory, we proceed to estimate the originally proposed model which has the benefit that it is easily recognisable as an extension of the seminal Bass and Mansfield models. The implication is that we cannot be certain that any evidence of an international effect is not due to spurious correlation. In light of the estimation results, we are somewhat less concerned since the general lack of robustness indicates that there is at least no strong bias in any one direction, if not towards a zero coefficient on the international effect.

3.4.2 Results

The world diffusion model provides a benchmark against which the country time series results can be evaluated. Estimates of the world Bass and Mansfield models for the periods 1837-1913 and 1837-1938 are presented in Table 3.3. The maximised log likelihood (a measure of fit which can be used when the dependent variables are the same) is slightly higher in the Bass model however the Bass parameter γ is not significantly different from zero. Estimates of the “endogenous” diffusion speed parameter β are similar in all four specifications: the parameter is positive and significant with a magnitude of about 0.13. Both models appear to fit better over the short than the long period: standard errors and the residual sum of squares are smaller. This raises the concern that the First World War years may represent a structural break or other discontinuity in the diffusion process. The point estimate of β is affected however the difference is not statistically significant. Estimates of the

β parameter are robust across the two time periods and the two models. As the benchmark we take the estimate of 0.11 from the Mansfield model over the short period.

Table 3.3 World diffusion estimates

Model	Period	N	β	se(β)	t(β)	γ	se(γ)	t(γ)	RSS	log L
B	1837-1913	76	0.13	0.01	9.0	0.00	0.00	-1.5	0.01	216.7
B	1837-1938	101	0.15	0.03	5.6	-0.01	0.01	-1.2	0.08	219.8
M	1837-1913	76	0.11	0.01	11.6				0.02	215.6
M	1837-1938	101	0.12	0.02	6.8				0.08	219.0

Notes: Model B is the Bass model (equation (3.6)) of world diffusion and Model M is the Mansfield model. N is the number of observations. RSS is the residual sum of squares. log L is the maximised log likelihood.

We now turn to discuss the results for individual countries. As is the case for international diffusion, estimates of the Bass coefficient γ are generally close to zero. We begin with the few cases in which this coefficient estimate is statistically significant. The results are presented in Table 3.4. The estimates are not very accurate and do not have the same sign. The negative sign obtained for Belgium and Finland is not consistent with the theoretical model because γ represents an information flow which theoretically cannot have a negative effect on the diffusion of a (superior) technology. Our main concern is the international effect, the existence of which is supported here in the sense that estimates are significantly different from zero in all cases except one. This exception is Belgium for which the model provides a very poor fit as indicated by high standard errors (especially for α), high RSS and insignificant β estimates. The poor fit is probably explained by the very fast pace of diffusion in Belgium where S(t)/T(t) had reached 90 per cent already in 1881 (Figure 3.2). The remaining estimates of α are not very encouraging either. For Finland and the United States the estimate of α is positive as hypothesised with a magnitude that is close to the β estimate in the benchmark model. However, the negative estimates of β make the international effect difficult to interpret because theoretically all of the parameters α, β and γ should be positive but this is not the case in any of the regressions. The Austrian results are peculiar because the magnitude of both the α and β estimates is very high. To sum up the general Bass model fits the data poorly: the Bass parameter γ is insignificant in most countries and where it is significant, other parameter estimates are not consistent with theory.

Table 3.4 Estimates of the extended Bass model (significant γ)

Country	Period	N	β	se(β)	t(β)	α	se(α)	t(α)	γ	se(γ)	t(γ)	RSS	log L
Austria	1837-1912	75	1.06	0.16	6.8	-0.91	0.16	-5.6	0.03	0.01	3.7	0.04	172.1
Belgium	1837-1931	94	0.27	0.22	1.2	0.91	0.51	1.8	-0.09	0.04	-2.4	0.73	94.9
Belgium	1837-1913	76	0.29	0.24	1.2	0.83	0.57	1.5	-0.08	0.04	-2.0	0.67	72.1
Finland	1873-1913	40	-0.38	0.19	-2.0	0.13	0.06	2.2	-0.02	0.01	-1.9	0.00	135.3
United States	1810-1913	103	-0.06	0.04	-1.4	0.14	0.04	3.8	0.01	0.00	1.9	0.02	287.1
United States	1837-1913	76	-0.19	0.08	-2.4	0.23	0.06	3.8	0.02	0.01	2.5	0.02	203.9

Notes: N is the number of observations. RSS is the residual sum of squares. log L is the maximised log likelihood.

The extended Mansfield model does considerably better than the extended Bass. Results of the extended Mansfield model, the original Mansfield model and the fully coupled model are presented in Table 3.5 to Table 3.7. The classic Bass model is not discussed because the parameter γ is not statistically different from zero except in the Austrian data.[36] Although all models were estimated over both time periods we present and discuss one period only. Generally this is the short period because the models tend to fit this data better. A short description of the choice of period is given in the Notes to each table.

We begin with the countries in which the estimates of the international effect (α) are consistent with the hypothesis i.e. the sign is positive (Table 3.5). In the extended

[36] The estimate was -0.010 with a standard error of 0.006.

model, the international effect is only significant at 5 per cent in France and the United States. At 0.1 and 0.11 the point estimates are close to the β estimate in the benchmark model but the 95 per cent confidence intervals are wide: (0.01, 0.19) for France and (0.04, 0.19) for the United States. Compared to the extended Bass model the point estimate appears to be considerably smaller in the United States data but the difference is not statistically significant. Nevertheless the difference suggests that if we exclude γ when it is in fact significant, any bias in estimates of α is likely to be negative. More problematic is the sign of the β estimates. In all cases, β is positive and significant in the Mansfield model but negative and insignificant in the extended model. We interpret this as evidence that the international effect is not working as we expected: rather than an additional explanatory variable, the international effect appears to dominate the domestic diffusion effect. In theoretical terms this means that when information from abroad is taken into account, information from domestic adopters does not have any value.

This raises the question of whether the fully coupled model is in fact the best representation. Indeed, it fits the data as well as the other two models and even slightly better than the Mansfield model. RSS and log L values are nearly equal for all models, and the point estimates of β and γ in the Mansfield and fully coupled models respectively are very close (except in Finnish data). This clearly suggests that the location of past diffusion does not matter: whether at home or abroad, the information value of diffusion is the same. That is, a one percentage point ceteris paribus increase in the steam- and motor ship share, whether at home or abroad, increases domestic diffusion by about 0.1 percentage points (as a share of total domestic tonnage). None of the models fits well to data from Australia and New Zealand which may be explained by the fact that the first observed values of

S(i,t)/T(i,t) are relatively high, at 18.3 and 22.5 per cent respectively.[37] We conclude that when the international effect is positive it is so strong that it dominates the (similarly positive) effect of past domestic diffusion on future diffusion.

Table 3.5 Country diffusion with a positive international effect

Country	Model	Period	N	β	se(β)	t(β)	α	se(α)	t(α)	RSS	log L
Australia	M	1876-1925	49	0.16	0.05	3.3				0.23	61.4
Australia	E	1876-1925	49	-0.34	0.57	-0.6	0.43	0.50	0.9	0.23	61.8
Australia	F	1876-1925	49				0.14	0.04	3.3	0.23	61.6
Finland	M	1873-1913	40	0.06	0.01	4.3				0.00	132.2
Finland	E	1873-1913	40	-0.05	0.07	-0.7	0.02	0.01	1.6	0.00	133.5
Finland	F	1873-1913	40				0.01	0.00	4.7	0.00	133.2
France	M	1838-1913	75	0.07	0.01	5.7				0.03	188.0
France	E	1838-1913	75	-0.04	0.06	-0.7	0.10	0.05	2.1	0.03	190.2
France	F	1838-1913	75				0.07	0.01	6.2	0.03	189.9
New Zealand	M	1870-1939	69	0.17	0.07	2.6				0.64	63.3
New Zealand	E	1870-1939	69	-0.64	0.56	-1.2	0.73	0.50	1.5	0.62	64.4
New Zealand	F	1870-1939	69				0.16	0.06	2.8	0.64	63.7
United States	M	1837-1913	76	0.08	0.01	7.1				0.03	193.5
United States	E	1837-1913	76	-0.01	0.03	-0.4	0.11	0.04	3.0	0.02	200.9
United States	F	1837-1913	76				0.09	0.01	8.2	0.02	200.9

Notes: Model M is the Mansfield model, E is the extended Mansfield model, and F is the fully coupled model. N is the number of observations. RSS is the residual sum of squares. log L is the maximised log likelihood. France and the United States: the models fit better to the short period (higher log L and t-values, better diagnostic test results[38]). Australia and New Zealand: long period with more data is preferred over short period with equally good (poor) fit. Finland: long period is inappropriate because there is a change in measurement (of tonnages) after 1919.

We now turn to the countries where the international effect is negative (Table 3.6). Immediately we see that in contrast to Table 3.5 there are no negative β estimates and all estimates are also significant at 5 per cent. The model fits generally well as there are no high RSS values. The international effect is significant at 5 per cent in Austria, Germany and Sweden; at 10 per cent in Italy; and insignificant in the Netherlands. The parameter is not very accurately estimated but tends to be less

[37] There was also a decline in diffusion during the First World War which contradicts the model assumptions and could explain the bad fit in the long period; however the fit is not any better for the short period.
[38] In this case, tests for autocorrelation and autoregressive conditional heteroskedasticity indicated problems in the long period but not in the short period.

than 0.1 in absolute value. Estimates of β are significantly higher in the extended model than the benchmark value 0.11 but not significantly different from the estimates in the simple Mansfield model with the exception of Austria.[39] The Austrian estimates stand out because both α and β are significantly larger in absolute value than in the other countries. This suggests diffusion in Austria reacts more strongly to changes in diffusion both home (positive effect) and abroad (negative effect). In the fully coupled model the estimates of α vary so that the benchmark estimate 0.11 is not included in all confidence intervals. The point estimate tends to be smaller than in the Mansfield model, and the Mansfield model also seems to fit the data somewhat better.

The results suggest that domestic diffusion has a significant positive effect on future diffusion while the use of the technology elsewhere has a negative effect. The magnitude of the international effect is roughly half of that of the domestic effect, which suggests that domestic information is of greater value than information from abroad. We conclude that the extended model appears to fit the data somewhat better when the international effect has a negative sign compared to when the sign is positive. However, the negative sign calls into question the theoretical explanation given to the international effect, that is, it is inconsistent with the argument that there is an information flow from abroad which has a positive effect on domestic diffusion. Indeed, we develop in the next chapter an alternative theoretical explanation in which the expected sign of the international effect is negative. Qualitatively, the results in Table 3.6 are the same as in Chapter 5 where we also find that the domestic effect is positive and the international effect is negative.

---

[39] Significance refers to 95 per cent confidence intervals. For example, confidence intervals for β in the extended Mansfield model do not contain the value 0.11.

Table 3.6 Country diffusion with a negative international effect

Country	Model	Period	N	β	se(β)	t(β)	α	se(α)	t(α)	RSS	log L
Austria	M	1837-1912	75	0.16	0.02	6.5				0.07	156.9
Austria	E	1837-1912	75	0.61	0.11	5.8	-0.40	0.09	-4.4	0.05	165.6
Austria	F	1837-1912	75				0.12	0.02	5.2	0.08	151.6
Germany	M	1850-1913	63	0.20	0.02	12.2				0.03	155.8
Germany	E	1850-1913	63	0.30	0.05	6.3	-0.08	0.04	-2.2	0.02	158.1
Germany	F	1850-1913	63				0.14	0.02	8.8	0.04	142.6
Italy	M	1862-1913	51	0.13	0.02	8.9				0.02	133.9
Italy	E*	1862-1913	51	0.20	0.04	4.7	-0.04	0.02	-1.7	0.01	135.2
Italy	F*	1862-1913	51				0.06	0.01	6.6	0.02	125.8
Netherlands	M	1846-1913	67	0.19	0.03	6.7				0.07	134.3
Netherlands	E	1846-1913	67	0.27	0.07	3.9	-0.07	0.06	-1.2	0.07	135.1
Netherlands	F	1846-1913	67				0.13	0.03	5.1	0.09	128.2
Sweden	M	1865-1913	48	0.15	0.02	8.5				0.02	116.2
Sweden	E	1865-1913	48	0.29	0.07	4.2	-0.10	0.04	-2.1	0.02	118.4
Sweden	F	1865-1913	48				0.09	0.01	6.9	0.03	110.6

Notes: Model M is the Mansfield model, E is the extended Mansfield model, and F is the fully coupled model. N is the number of observations. RSS is the residual sum of squares. log L is the maximised log likelihood. *Models E* and F* use W_{t-1}/TW_{t-1} instead of $(W_{t-1}-S_{t-1})/(TW_{t-1}-T_{t-1})$. The models fit better to the short than the long period in each case. Netherlands: poor fit to the long period (RSS around 1.0) is probably explained by the lack of change in diffusion between 1921 and 1938. Italy: fit is better for the short period (RSS smaller, log L higher, less diagnostic problems[40]). Germany and Sweden: t-ratios are higher in the short period. There is no data for Austria after 1912.

Finally, there are five countries for which the international effect is not statistically significant (Table 3.7). There is no econometric reason why these countries are any different from the others already discussed; insignificance of α is not due to particularly high standard errors, for example. Estimates of β are close to the benchmark 0.11 except in Belgian data, to which all models fit poorly. Results of the basic and extended Mansfield models hardly differ except that the standard error of β is considerably higher in the extended model. There does not seem to be any benefit in including γ in the model. Surprisingly in this light, the fully coupled model nearly fits the data as well as the Mansfield model and even better in the case of the

[40] In this case, tests for autocorrelation and autoregressive conditional heteroskedasticity indicated problems in the long but not in the short period, for models M and E*.

United Kingdom (i.e. log L is higher in the fully coupled model). Estimates of α in the fully coupled model tend to be smaller than estimates of β in the Mansfield model although the difference is only significant in Norwegian data. This can be interpreted so that foreign information is not as valuable for diffusion as domestic information.

Table 3.7 Country diffusion with no international effect

Country	Model	Period	N	β	se(β)	t(β)	α	se(α)	t(α)	RSS	log L
Belgium	M	1837-1931	94	0.43	0.10	4.2				0.78	92.0
Belgium	E	1837-1931	94	0.44	0.21	2.0	-0.02	0.33	-0.1	0.78	92.0
Belgium	F	1837-1931	94				0.58	0.16	3.6	0.81	89.9
Canada	M	1892-1939	47	0.08	0.02	4.9				0.03	106.7
Canada	E*	1892-1939	47	0.08	0.08	1.0	0.00	0.04	0.0	0.03	106.7
Canada	F*	1892-1939	47				0.04	0.01	4.7	0.03	106.2
Denmark	M	1844-1913	69	0.15	0.02	9.4				0.03	169.4
Denmark	E	1844-1913	69	0.17	0.06	2.8	-0.02	0.05	-0.4	0.03	169.5
Denmark	F	1844-1913	69				0.11	0.01	8.5	0.03	165.6
Norway	M	1866-1913	47	0.13	0.01	12.2				0.01	144.7
Norway	E	1866-1913	47	0.14	0.02	6.4	0.00	0.01	-0.2	0.01	144.7
Norway	F	1866-1913	47				0.05	0.01	7.5	0.01	129.6
United Kingdom	M	1837-1913	76	0.13	0.02	7.8				0.04	176.9
United Kingdom	E	1837-1913	76	0.12	0.07	1.7	0.02	0.09	0.3	0.04	179.7
United Kingdom	F	1837-1913	76				0.18	0.02	7.5	0.04	178.3

Notes: Model M is the Mansfield model, E is the extended Mansfield model, and F is the fully coupled model. *Models E* and F* use W_{t-1}/TW_{t-1} instead of $(W_{t-1}-S_{t-1})/(TW_{t-1}-T_{t-1})$. N is the number of observations. RSS is the residual sum of squares and log L is the maximised log likelihood. Belgium: the models fit poorly to both periods so the longer period is presented. Canada: only a long period is estimated. Denmark: models fit poorly to the long period (RSS 0.63, log L around 103) possibly because of a decline in diffusion during the First World War. Norway and United Kingdom: fit is poorer for the long period (log L and t-values are smaller).

To conclude, the results provide some support for the existence of an international effect however the direction of this effect is not robust since there is an equal number of positive, negative and zero estimates. Most interestingly, the fit of the extended model is best when the α estimate is negative, which is inconsistent with the hypothesis that information from abroad encourages further use. On the other hand when α is positive as expected, the domestic effect parameter β is negative, which is again inconsistent with theory because this effect is also based on an information-spreading argument. In our data, there is no country for which estimates of both α and β are positive. Instead, we find evidence that domestic and international diffusion always have contradictory effects, with the international effect sometimes dominating.

The estimate of β in the benchmark model is of a magnitude that we would expect. Most of the other estimates of β contain the benchmark within the 95 per cent confidence interval. Estimates are generally less precise in the general model than in either of the models which have just one parameter. The length or start date of the study period do not appear to be related to the estimated magnitudes. Quite naturally as the two biggest contributors to world diffusion, results for United Kingdom and United States are close to the benchmark. All models fit poorly to data from Australia, New Zealand and Belgium. This can be explained by the nature of the diffusion series: the first two data series begin with a very high first observation and the logistic curve is known to fit poorly if early data is missing, which may be the explanation here. The Belgian estimates are poor probably because diffusion reached a high value very early on. After this date the models have little explanatory power over what are only very small annual changes in S(i,t)/T(i,t), some of which are negative.

3.5 Conclusion

The objective of this chapter was twofold: first, to make a theoretical case for the international effect in an epidemic model of diffusion and second, to test the resulting hypothesis using data on steam- and motor ship diffusion. Our theoretical argument is that the extent of use abroad is a source of information about the technology, its characteristics and profitability. The greater the use of new technology elsewhere, the more information about the technology is available to potential domestic users and therefore the lesser the degree of uncertainty about the benefits of adoption. The hypothesis is therefore that international diffusion has a positive effect on the domestic extent of use.

We introduced international diffusion to the Bass (1969) and Mansfield (1961) models as an additive effect similar to the γ parameter of the Bass model. World diffusion is measured as the simple average of steam- and motor ship diffusion in 13 countries, from which each country's own contribution is subtracted so that the explanatory variable is the extent of use elsewhere. Using non-linear least squares on country time-series for two periods of different length, we found some evidence that a world effect exists. The sign of the effect is not robust: there is an equal number of positive, negative and insignificant estimates. The time-series properties of the model may explain this to some degree. However, it is clear that the negative coefficient estimates are inconsistent with the theoretical argument that international diffusion matters because use abroad is a source of information.

We need to examine then whether a theoretical explanation can be made for a negative international effect. It seems to us that this is difficult given the focus on information-spreading unless we relax the assumption that all information has the same value. The history of shipping reveals that technological progress and market segmentation interacted. By this we mean that sailing ships were first replaced by steamships on shorter routes and the transport of passengers and valuable cargo but as steam technology (in particular fuel efficiency) improved, steamships were adopted on ever longer routes. This suggests a hypothesis which is consistent with the lack of robustness in our results, namely that ship owners are primarily interested in information which comes from others operating in the same market. If so, then information from the early steamships which are not yet competitive on all routes is of low value to those sailing ship owners who operate exclusively in the long-haul market. In this case, the *lack* of international diffusion acts as a negative signal which may be stronger than the positive sign given by the adoption of steam- and motor ships in markets that are not relevant. If we put together all diffusion into one international measure estimation results will be muddled because all information is given the same weight.

The challenges posed by this argument are considerable. The study should begin with a definition of markets, say according to length of journey and the nature of the cargo. This argument leads towards a desire for micro-level data i.e. information on the level of the ship owner. Since we have only country-level data, we could instead attempt to define for each country the proportion of their fleet that operates in each market. The international diffusion measure would then reflect the degree to which diffusion elsewhere is based in the markets in which the domestic fleet operates. Models developed in the spatial econometrics literature may be of use here. There is a way forward then even with country-level data, however in our view the approach just outlined would require such detail that the challenge is not proportional to the objective which is to provide a first attempt to extend a closed-economy epidemic model to include a link with the extent of diffusion elsewhere. We have not intended to build a model particularly of steam- and motor ship diffusion but rather to use this as an example or context in which the international diffusion hypothesis can be tested.

The lack of robustness in our results has two implications in terms of the estimation method. First, panel data methods may provide more robust results than individual time series because in a panel, information about cross-country variation in diffusion can be used in addition to the time-series information which we have relied on here. On the other hand, a drawback of standard panel methods is that parameters are assumed to common, that is, it is assumed that the true values of α, β and γ are the same for all countries. If this assumption is false and parameters are in fact country-specific – for example, the international effect is only positive for a subset of countries – then panel estimates will be inconsistent. Second, the lack of robustness may also be due to nonstationarity. We tested the time series properties of variables, allowing for structural breaks in the time series, and found that nonstationarity cannot be rejected in the variables of the classic epidemic models. This means that it is possible that the results are due to spurious correlation. The question is then how to develop an epidemic diffusion model which does not have this problem. We experimented with a logistic transformation of the diffusion measure and found that nonstationarity can be rejected in most of the transformed series. However, the results were not robust enough to be satisfactory. It seems to us that another important direction in which the study here could be extended is to take the time-series properties of the model as the main objective. This would potentially make a considerable contribution to the epidemic literature.

We conclude our discussion of epidemic models here. In the next two chapters we investigate the relationship between international and domestic diffusion in an alternative theoretical framework, so-called decision-theoretic models. The international effect is given a wholly different role in the diffusion process as information-spreading is not the driving force of diffusion in these models. The result is a hypothesis that the international effect is negative. We begin by developing a detailed model of how in a world of Cournot competition adoption of the new technology by foreign competitors changes the benefit of adoption. In Chapter 5 the hypothesis is tested using panel data on the diffusion of the basic oxygen furnace, also a process technology.

Appendix to Chapter 3: Time-series properties

The first two sections of this appendix explain the concept of stationarity and the testing procedure. Section A.3 presents the results referred to in Chapter 3. In the final section we show that tests are much more successful (in rejecting nonstationarity) when applied to the logistic transformation of the diffusion measure.

A.1 Stationarity

A time-series process is said to be strictly stationary if its properties are unaffected by a change of time origin (Verbeek 2000:228). We are concerned with weak stationarity, which implies that the means, variances and covariances of the series are independent of time. Thus a stationary series fluctuates around its constant mean with a constant finite variance. In contrast, non-stationary series have time-dependent means, variances or autocovariances, and graphically they typically "wander off" instead of returning periodically to the long-run mean. An example of a stationary series is

$$Y(t) = \delta + \mu Y(t-1) + u(t) \quad where\ |\mu| < 1\ and\ u(t) \sim \mathrm{IN}(0, \sigma^2).$$

The mean is $\delta/(1-\mu)$ and the variance is $\sigma^2/(1-\mu^2)$. An example of a non-stationary series is a random walk with drift

$$Y(t) = \delta + Y(t-1) + u(t) \quad u(t) \sim \mathrm{IN}(0, \sigma^2).$$

Here the autoregressive coefficient μ is unity and the unconditional variance does not exist, i.e. it is infinity. Thus stationarity is violated. In this case the non-stationarity can be eliminated by taking first differences. The series is said to have one unit root or to be integrated of order 1, denoted I(1).

The main reason for testing the degree of stationarity of economic variables is concern for the "spurious regression" problem. Let Y(t) and X(t) be two variables generated by independent pure random walks:

$$Y(t) = Y(t-1) + u(1,t) \quad u(1,t) \sim IID(0, \sigma_1^2)$$
$$X(t) = X(t-1) + u(2,t) \quad u(2,t) \sim IID(0, \sigma_2^2)$$

where u(1,t) and u(2,t) are mutually independent. Granger and Newbold (1974) showed that in the OLS regression

$$Y(t) = \delta + \mu X(t) + \nu(t)$$

the t-ratio on μ is likely to be significant, despite the lack of a causal relationship between Y(t) and X(t). To rule out the possibility of spurious regression we test for the degree of stationarity in our variables, before proceeding with least-squares estimation.

A.2 Testing for stationarity

The standard test for stationarity is the Dickey-Fuller test and its extension the augmented Dickey-Fuller (ADF) test. In this approach the null hypothesis is a unit root and the alternative is stationarity. Sometimes non-stationarity is caused not by a unit root but by a deterministic linear time trend. In this case the appropriate alternative hypothesis is

$$Y(t) = \delta + \phi t + \mu Y(t-1) + \varepsilon(t) \quad |\mu| < 1, \gamma \neq 0 .$$

The process is trend stationary: the mean (δ+φt)/(1-μ) is time-dependent, however this non-stationarity can be removed simply by including t as an additional variable in the model; or by regressing Y(t) against a constant and t, and then considering the residuals of this regression (Verbeek 2000:239-40).

To test the null of a unit root in Y(t) against the alternative of (trend) stationarity, we run the ADF regression

$$\Delta Y(t) = \delta + \phi t + (\mu - 1)Y(t-1) + \sum_{j=1}^{n} \Delta Y(t-j) + u(t) \,. \tag{1}$$

The specification of the lag length n assumes that u(t) is white noise. Various selection criteria have been proposed for choosing the lag length; we begin with 15 lags and use primarily the t-statistic on the longest lag. If the lag is significant at 5 per cent, that determines the lag length. If the longest lag is significant at 10 per cent, we use the Akaike information criterion (AIC) and the equation standard error to determine whether the lag can be dropped. When the lag length has been chosen the t-ratio on (μ-1) is used to test the null hypothesis. Dickey and Fuller (1979) showed that this does not have a standard t-distribution because of the non-stationarity of the process. We use critical values from MacKinnon (1991). The critical values depend on whether a constant δ, or a constant and a time trend φt, are included in (1). The intercept δ is usually included, because the alternative δ=0 would impose a zero mean in the stationary series. We will initially include the term φt in all ADF regressions however if a unit root is not rejected we test for the significance of the trend using critical values from Dickey and Fuller (1981, Tables V and VI). To include a trend when it is not significant reduces the power of the test.

Following standard practice if the null of a unit root is not rejected we then test for stationarity in the first differences, i.e. ΔY(t). The procedure is exactly as above except that the dependent variable is the second difference, Δ^2Y(t). The test statistic is t(μ-1) from

$$\Delta^2 Y(t) = \delta + \phi t + (\mu - 1)\Delta Y(t-1) + \sum_{j=1}^{n} \Delta^2 Y(t-j) + u(t) \,.$$

If the ADF test for Y(t) does not reject a unit root, but does reject it in first differences, then we conclude that Y(t) has one unit root. If both Y(t) and ΔY(t) are nonstationary but Δ^2Y(t) is stationary, then Y(t) has two unit roots.

The ADF test has low power if the time series is stationary but very close to a unit root, or if the series has a nonlinear trend or structural breaks. Perron (1989) developed a dummy variable unit root test which is appropriate when the time series has a single structural break. In Perron's test the break affects either the intercept, the slope of the trend function, or both. The critical values depend on the ratio of the pre-break sample size to total sample size, denoted by λ. Let T_B denote the date of the structural break. Define a dummy variable $DL(T_B)$

$$DL(T_B) = \begin{matrix} 0, if & t \leq T_B \\ 1, if & t > T_B \end{matrix} .$$

Let $DP(T_B)$ be an impulse dummy that takes the value 1 in year T_B+1 and the value 0 in every other period. Perron (1989) presents 3 alternative null hypotheses. Consider first the null that Y(t) is a unit root with a single break in the intercept. The alternative hypothesis is trend stationarity with a single change in the intercept. To test the null we run the regression

$$\Delta Y(i,t) = \delta + \phi t + (\mu - 1)Y(i,t-1) + \delta_1 DL(T_B) + \delta_2 DP(T_B) + \sum_{j=1}^{n} c(j)\Delta Y(i,t-j) + u(i,t) . \qquad (2)$$

Perron (1989, Table IVB) presents critical values for the t-statistic on (μ-1). As in the simple ADF test the lag length n is chosen so that ε(i,t) is IID. The coefficient δ_2 captures the immediate impact of the shock while the permanent change in the intercept is δ_1. Second, let the null be a unit root with a break in the trend only. Define DT (T_B) as a dummy

$$DT(T_B) = \begin{matrix} 0, if & t \leq T_B \\ t, if & t > T_B \end{matrix} .$$

We test the null of a unit root with

$$\Delta Y(i,t) = \delta + \phi t + (\mu - 1)Y(i,t-1) + \delta_3 DT(T_B) + \sum_{j=1}^{n} c(j)\Delta Y(i,t-j) + u(i,t) . \qquad (3)$$

Again the lag length n is chosen so that the residual u(i,t) is white noise. Critical values for t(μ-1) are found in Perron (1989, Table VB). As a third possible null hypothesis, Perron (1989) considers a unit root with a break in both the intercept and trend at time T_B. The alternative is trend stationarity, with a structural break. The appropriate regression is

$$\Delta Y(i,t) = \delta + \phi t + (\mu - 1)Y(i,t-1) + \delta_1 DL(T_B) + \delta_2 DP(T_B) + \delta_3 t * DL(T_B) + \sum_{j=1}^{n} c(j)\Delta Y(i,t-j) + u(i,t) \quad (4)$$

with the lag length n chosen so that the residual u(i,t) is white noise. The interaction term $t*DL(T_B)$ captures the change in the time trend at T_B. Critical values are provided for t(μ-1) in Perron (1989, Table VIB).

A.3 Results

We have used the following methodology to conduct the unit root tests. First we test for a unit root using the simple ADF test without structural breaks (1). The trend term φt is included unless it is insignificant (see above) and a unit root is rejected in the ADF test with intercept only (note that the trend term is always included in Perron's test.) If a unit root is rejected in the simple ADF test, this is the result reported below. If not, and we have a suggestion for a T_B, i.e. a possible structural break, we run Perron's test with a break in both intercept and trend, i.e. (4). Depending on the significance of the break terms we also run either (2) or (3). In each stage we test for the appropriate lag length n although this does not typically vary between the different specifications of Perron's test. If Perron's test rejects or is considerably closer to rejecting a unit root than the simple ADF test, we report the result of Perron's test below. All tests were conducted for the same (two) study periods for which the diffusion models were estimated. Here we report results for the period for which estimates are reported in Section 3.4.2 unless a unit root was only rejected for the other period, in which case results are given for both periods.

Ideally we would be able to refer to external sources to determine the date of the structural break T_B in each country. However, this choice is only clear for Finland where there was a change in measurement in 1919.[41] We have two reasons for considering structural breaks. The first is practical: to increase the power of the ADF test which is known to be low in the presence of nonlinearities. This concern suggests using an F-test to pick the value of T_B that gives the strongest support for a structural break. This motivation is rather opportunistic and fortunately we also have a second reason to consider structural breaks: that is, true structural breaks are likely to have occurred given the sheer length of the time series. Breaks may have been country-specific, in which case T_B must be determined separately for each country, or international, in which case it may be justified to fix T_B to the same value for all countries, say the start of the First World War. We want to allow for differences in T_B across countries because countries may have been more or less vulnerable to international shocks depending on the stage of their diffusion process, size of the commercial fleet, geographical location, etc. Diffusion in early adopting countries, say, could have been more vulnerable to an international shock in 1889 than in 1919. Also, we want to allow for time lags in the effects of any international shocks by letting T_B be country-specific.

Our chosen approach is to look for possible dates for structural breaks using the plot of S(i,t)/T(i,t) with special consideration given to war years and then to apply F-tests to find a year most appropriate for each country. To illustrate the F-test approach,

[41] Ships below a certain tonnage were excluded from the statistics from 1919 onwards. This dramatically decreased the number of sail ships since this was the most common type among small vessels.

denote the start of the study period by t=0 and the final period by t=T. We divide the study period into two sub-periods around the break, at $t=T_B$. The following regression is run three times, first for the whole period t=0 to t=T, and then for the two sub-periods: t=0 to $t=T_B$, and $t=T_B+1$ to t=T:

$$Y(i,t) = \alpha + \gamma t + \varepsilon(t).$$

The residual sum of squares (RSS) is obtained from each of these 3 regressions. The test statistic is calculated as

$$F = \frac{R_1 / k}{R_2 / (N_1 + N_2 - 2k)},$$

where k is 2, the number of parameters; R_1 is the RSS from the full period; and R_2 is the sum of the RSS's from the two sub-periods. The test is run for different choices of $t=T_B$ and the one with the highest test statistic is chosen as the date of the structural break. If the logistic transformation clearly failed to eliminate the nonlinear trend in the data, we considered T_B chosen by an F-test from the regression

$$Y(i,t) = a + bt + dt^2 + v(t).$$

In this case k=3 in the test statistic. Since the purpose of the structural break is to increase the power of the ADF test which does not allow a nonlinear trend, we mostly chose T_B from the former regression without a quadratic trend.

Consider the linear version of the general model (equation (3.11)):

$$\frac{S_{i,t} - S_{i,t-1}}{T_{i,t-1}} = \gamma + (\beta - \gamma)\frac{S_{i,t-1}}{T_{i,t-1}} - \beta \frac{S_{i,t-1}^2}{T_{i,t-1}^2} + \alpha \frac{W_{t-1} - S_{i,t-1}}{TW_{t-1} - T_{i,t-1}} - \alpha \frac{S_{i,t-1}\left(W_{t-1} - S_{i,t-1}\right)}{T_{i,t-1}\left(TW_{t-1} - T_{i,t-1}\right)} + \varepsilon_{i,t}$$

In the Bass and Mansfield models α=0 and there are only two right-hand side variables. Unit root test results for these and the left-hand side variable are presented in Table A1. We find that the left-hand side variable is stationary (Panel A) but the right-hand side variables are not with a few exceptions (Panels B and C).

In the general model there are two additional variables that should be tested for stationarity: from (3.11) these are

$$\frac{W_{t-1} - S_{i,t-1}}{TW_{t-1} - T_{i,t-1}} \quad \text{and} \quad \frac{S_{i,t-1}\left(W_{t-1} - S_{i,t-1}\right)}{T_{i,t-1}\left(TW_{t-1} - T_{i,t-1}\right)}.$$

Results of stationarity tests for these variables are presented in Table A2. Nonstationarity cannot be rejected in the first variable and it is only rejected in the second variable for a few countries. We conclude that the right-hand side variables of the extended model, and therefore also of the extended Mansfield, original Bass and original Mansfield models appear to be nonstationary. A unit root can only be rejected in the left-hand side variable.

Table A1. Unit root tests for the Bass and Mansfield model variables

Panel A: ∆S(t)/T(t-1)

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	A		0	0.17	Y	-5.57
Austria	1837-1912	B	1894	7	-0.48	Y	-6.02
Belgium	1837-1931	A		4	0.42	5%	-2.98
Canada	1892-1939	A		0	0.41	Y	-4.28
Denmark	1844-1913	A		0	0.36	Y	-5.55
Finland	1873-1913	A		0	-0.06	Y	-6.16
France	1839-1939	A		2	0.36	Y	-6.52
France	1838-1913	A		9	0.33	N	-2.53
Germany	1850-1938	B	1913	11	-0.03	Y	-3.83
Germany	1850-1913	A		11	0.82	N	-1.41
Italy	1862-1925	A		7	-3.44	Y	-4.58
Italy	1862-1913	A		12	1.58	N	2.06
Netherlands	1846-1913	A		5	-0.41	Y	-4.71
New Zealand	1870-1939	A		0	-0.20	Y	-9.90
Norway	1866-1913	A		2	-0.05	Y	-6.15
Sweden	1865-1913	A		1	0.17	Y	-4.40
United Kingdom	1837-1913	A		0	-0.23	Y	-10.76
United States	1837-1939	A		1	0.63	Y	-5.08
United States	1837-1913	A		2	0.54	5%	-3.15
World	1837-1938	A		2	0.42	Y	-5.81
World	1837-1913	A		0	-0.05	Y	-8.91

Notes: "Model" indicates the specific unit root test that was used: A allows no structural break; B allows a single break in the intercept; C a break in the slope; D a break in both intercept and slope. T_B is the date of the structural break. μ-ADF is the estimate of the autoregressive coefficient. Column "reject?" indicates whether a unit root is rejected at 1 per cent (Y), at 5 per cent (5%) or not (N). n is the number of lags. t-ADF is the t-ratio on (μ-1).

Table A1. continued

Panel B: S(t-1) / T(t-1)

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	B	1919	0	0.82	N	-1.93
Austria	1837-1912	A		5	0.98	N	-1.78
Belgium	1837-1931	A		13	0.95	N	-1.97
Canada	1892-1939	D	1921	0	0.59	N	-3.33
Denmark	1844-1913	D	1871	1	0.86	N	-2.89
Finland	1873-1939	D	1920	0	0.34	Y	-6.94
Finland	1873-1913	A		0	0.64	N	-2.66
France	1838-1913	A		8	0.92	N	-2.79
Germany	1850-1913	A		4	0.96	N	-2.32
Italy	1862-1913	A		10	0.91	N	-1.54
Netherlands	1846-1913	A		13	0.88	5%	-4.00
New Zealand	1870-1939	C	1913	8	0.40	Y	-5.01
Norway	1866-1913	A		0	0.99	N	-0.48
Sweden	1865-1913	A		13	0.81	N	-2.97
United Kingdom	1837-1913	A		1	0.94	N	-2.60
United States	1837-1913	A		11	0.93	N	-1.25
World	1837-1938	A		6	0.97	N	-1.70
World	1837-1913	D	1865	0	0.73	N	-3.57

Panel C: [S(t-1) / T(t-1)]^2

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	B	1919	0	0.55	N	-3.62
Austria	1837-1912	D	1878	9	0.88	N	-3.81
Belgium	1837-1931	A		5	0.96	N	-2.21
Canada	1892-1939	D	1921	0	0.75	N	-2.47
Denmark	1844-1913	D	1871	6	0.96	N	-1.59
Finland	1873-1939	D	1920	5	-0.21	Y	-7.09
Finland	1873-1913	A		2	0.72	N	-2.40
France	1838-1913	A		14	0.89	N	-2.12
Germany	1850-1913	A		6	0.96	5%	-3.87
Italy	1862-1913	A		19	2.39	N	2.17
Netherlands	1846-1913	A		13	0.81	Y	-4.57
New Zealand	1870-1939	B	1913	3	0.76	N	-3.24
Norway	1866-1913	A		15	2.31	N	2.58
Sweden	1865-1913	A		13	0.85	N	-2.32
United Kingdom	1837-1913	A		0	0.99	N	-0.83
United States	1837-1913	A		1	1.01	N	0.81
World	1837-1938	A		12	0.97	N	-3.07
World	1837-1913	D	1865	3	0.95	N	-2.35

Notes: “Model” indicates the specific unit root test that was used: A allows no structural break; B allows a single break in the intercept; C a break in the slope; D a break in both intercept and slope. T_B is the date of the structural break. μ-ADF is the estimate of the autoregressive coefficient. Column “reject?” indicates whether a unit root is rejected at 1 per cent (Y), at 5 per cent (5%) or not (N). n is the number of lags. t-ADF is the t-ratio on (μ-1).

Table A2. Unit root tests for the additional variables in the general model

Panel A: [W(t-1) - S(t-1)] / [TW(t-1) - T(t-1)]

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	A		3	0.95	N	-0.64
Austria	1837-1912	D	1865	1	0.70	N	-3.77
Belgium	1837-1931	A		6	0.96	N	-2.09
Denmark	1844-1913	A		0	0.97	N	-2.00
Finland	1873-1913	A		0	0.97	N	-2.13
France	1838-1913	A		0	0.97	N	-2.18
Germany	1850-1913	A		0	0.91	N	-3.00
Netherlands	1846-1913	A		0	0.97	N	-2.09
New Zealand	1870-1939	D	1913	1	0.79	N	-2.93
Norway	1866-1913	A		0	0.82	N	-1.96
Sweden	1865-1913	A		0	0.77	N	-2.61
United Kingdom	1837-1913	A		11	0.98	N	-0.78
United States	1837-1913	D	1864	1	0.74	N	-3.52

Panel B: [S(t-1) / T(t-1)] * [W(t-1) - S(t-1)] / [TW(t-1) - T(t-1)]

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	D	1919	0	0.69	N	-3.39
Austria	1837-1912	D	1878	12	0.85	N	-3.61
Belgium	1837-1931	A		13	0.92	N	-3.26
Canada	1892-1939	D	1921	0	0.69	N	-2.73
Denmark	1845-1939	B	1897	5	0.94	Y	-3.90
Denmark	1844-1913	D	1871	6	0.96	N	-1.84
Finland	1873-1939	D	1920	0	0.25	Y	-8.18
Finland	1873-1913	A		0	0.78	N	-2.79
France	1838-1913	A		6	0.97	N	-1.22
Germany	1850-1913	A		6	0.97	N	-2.75
Italy	1862-1913	A		16	1.42	N	2.61
Netherlands	1846-1913	A		13	0.87	Y	-4.96
New Zealand	1870-1939	A		1	0.91	N	-2.00
Norway	1866-1913	A		11	0.95	N	-1.03
Sweden	1865-1913	A		9	0.96	N	-1.56
United Kingdom	1837-1913	A		0	1.01	N	1.37
United States	1837-1913	D		0	0.98	N	-1.28

Notes: "Model" indicates the specific unit root test that was used: A allows no structural break; D allows a break in both intercept and slope. T_B is the date of the structural break. λ is the ratio of the pre-break sample to total sample size. μ-ADF is the estimate of the autoregressive coefficient. Column "reject?" indicates whether a unit root is rejected at 1 per cent (Y), at 5 per cent (5%) or not (N). n is the number of lags. t-ADF is the t-ratio on (μ-1). Canada and Italy are missing from Panel A because since they are not included in W(t) and TW(t) the ADF test on the variable [W(t)-S(i,t)]/[TW(t)-T(i,t)] is not relevant for them.

A.4 Logistic transformation of the diffusion measure

We have found evidence that nonstationarity is a problem in the Bass and Mansfield models. In particular, a unit root can generally be rejected in the left-hand side variable [S(i,t)-S(i,t-1)]/T(i,t) but not in the diffusion measure S(i,t)/T(i,t). In this section we discuss a possible approach to solving this problem namely applying the logistic transformation to S(i,t)/T(i,t). If the transformed variable is stationary, it provides a possible basis for a new type of epidemic model in which all variables have the desired time-series properties. We will first describe how the transformation is performed and then discuss the results of unit root tests on the transformed variable. We find that non-stationarity is now rejected more commonly however not for all countries. Because it remains a problem, the logistic transformation is not the basis for a new stationary model that we are looking for.

In Dickey-Fuller type unit root tests a key assumption of the alternative hypothesis is that the time trend of the stationary series is linear. This assumption contradicts the argument of the Bass and Mansfield models that diffusion data follows an S-shaped trend over time. However if this S-shape takes say the logistic function form then applying the logistic transformation should produce a transformed diffusion measure with a linear time trend.

Consider the steam- and motor ship tonnage S(i,t). The logistic transformation of this variable is given by

$$S'(i,t) = \ln \frac{S(i,t)}{S^*(i,t) - S(i,t)} . \tag{5}$$

Since we want to analyse the variable S(i,t)/T(i,t) we divide by T(i,t) and assume that S*(i,t)=T(i,t) as before. The transformation is now given by

$$S'(i,t) = \ln \frac{S(i,t)/T(i,t)}{1 - S(i,t)/T(i,t)} . \tag{6}$$

Similarly the logistic transformation of the world diffusion measure is given by

$$W'(i,t) = \ln \frac{W(t)/TW(t)}{1 - W(t)/TW(t)} \tag{7}$$

which we have obtained by assuming W*(t)=TW(t).

Plotting the transformed series for each of the 15 countries reveals a markedly more linear time trend, i.e. the S-shape is mostly no longer visible. We tested the transformed variables S'(i,t) and W'(t) for stationarity using either the ADF test or Perron's unit root test depending on whether there was an indication of a structural break. The results are presented in Table A3. We can reject non-stationarity without a structural break in the data for Canada, Finland, Italy and the Netherlands. Allowing for a break in the intercept and/or slope of the time trend, a unit root can be rejected for Austria, Belgium, Denmark, New Zealand and the world diffusion measure W'(i,t); and also in the long period for France, Germany, Norway and the United States. We cannot reject non-stationarity for Australia, Sweden or the United Kingdom. The first difference ΔS'(i,t) appears to be stationary which indicates that the data has at most one unit root.[42]

[42] A unit root in the first differences is rejected for all countries except in the short period for France, Germany and the world diffusion measure.

The fact that structural breaks have to be considered for non-stationarity to be rejected suggests that the logistic functional form does not fit the data well. Although a unit root can be rejected for most countries at least in the long period, we conclude that the logistic transformation is not the way forward and it is not worthwhile to further explore how it could be used to develop a new extended model with a world diffusion effect. Although the experiment does not have the desired result, it has served to demonstrate the scale of the challenge and the importance of considering the time series properties of models.

Table A3. Unit root tests for S'(i,t) and W'(t)

Country	Model	Period	T_B	μ-ADF	Reject?	n	t-ADF
Australia	B	1876-1925	1918	0.76	N	0	-2.60
Austria	B	1837-1912	1896	0.64	Y	7	-4.24
Belgium	C	1838-1931	1893	0.76	Y	6	-4.04
Canada	A	1892-1939		0.96	Y	0	-3.48
Denmark	B	1845-1913	1889	0.70	Y	6	-4.03
Finland	A	1873-1913		0.92	Y	4	-3.72
France	D	1838-1939	1897	0.82	Y	5	-4.25
France	A	1838-1913		0.99	N	8	-1.81
Germany	D	1850-1938	1902	0.62	Y	7	-4.42
Germany	A	1850-1913		0.66	N	7	-2.73
Italy	A	1862-1925		0.75	Y	1	-3.59
Italy	A	1862-1913		0.70	5%	1	-4.05
Netherlands	A	1846-1913		0.73	Y	3	-3.65
New Zealand	D	1870-1939	1917	0.20	Y	8	-5.17
Norway	D	1866-1938	1919	0.41	Y	7	-4.63
Norway	A	1866-1913		0.73	N	1	-3.50
Sweden	A	1865-1913		0.85	N	4	-1.47
United Kingdom	A	1837-1913		0.86	N	1	-2.97
United States	C	1837-1939	1901	0.67	Y	7	-4.03
United States	A	1837-1913		0.71	N	11	-2.07
World diffusion	D	1838-1938	1858	0.72	Y	1	-3.99
World diffusion	D	1838-1913	1858	0.04	Y	11	-4.86

Notes: "Model" indicates the specific unit root test that was used: A allows no structural break; B allows a single break in the intercept; C a break in the slope; D a break in both intercept and slope. T_B is the date of the structural break. μ-ADF is the estimate of the autoregressive coefficient. Column "reject?" indicates whether a unit root is rejected at 1 per cent (Y), at 5 per cent (5%) or not (N). n is the number of lags. t-ADF is the t-ratio on (μ-1). Belgium: a unit root is only rejected if the first observation is not included.

4 Cournot model of international diffusion

Equation Chapter 4 Section 1

4.1 Introduction

Our objective in this chapter is to give microeconomic foundations to the international effect, that is, to analyse how international diffusion affects the individual producer's adoption decision. We do this by specifying how the extent of use affects the costs and benefits of adopting today versus postponing adoption. The model is an extension of Reinganum (1981b) in which the key feature is Cournot competition in the output market. Demand is assumed to be linear and the price of technology is given, that is, exogenous to diffusion. A low-cost process technology is available and each producer decides when to adopt it. The price of output depends on the number of producers who use the new technology. As the extent of use increases, the price of output falls. The key feature of the model is a so-called stock effect which is a relationship between the extent of use and further use. As the price of output falls due to diffusion, adoption is less profitable in the sense that the gross benefit (per-period profit) of using the new technology relative to the old one falls. We propose that the effect of international diffusion on domestic diffusion can be modelled via this stock effect.

We extend the Reinganum model by allowing producers to be heterogeneous with respect to the costs of production so that the cost advantage of the new technology is producer-specific. While Reinganum (1981b) is a closed-economy model we interpret it in an international context by specifying that producers are located in different countries. This gives us the opportunity to analyse the relationship between domestic and international diffusion. Also, our model provides a framework to which institutional variables such as education and political institutions are easily introduced.

Although the theoretical discussion is intended to be general, throughout the chapter we refer to a specific example of diffusion, the basic oxygen furnace. Data on this technology is used in Chapter 5 to test the hypotheses that are developed here. The reference to a particular technology is also an aid to presentation and discussion of model assumptions and implications. A detailed discussion of the historical case is provided in the next chapter but the stylised story is the following. The basic oxygen furnace is a process technology used to produce crude steel. Before 1952, there are two main production technologies. The oldest is the Bessemer process but this has largely been replaced by the lower-cost open hearth furnace. A new method, the basic oxygen furnace, is now invented in Austria. Producers around the world are quickly informed of its superiority in terms of lower installation and production costs. However adoption is costly, and some producers continue to use open hearths for several decades. Only in 1992 does the diffusion process come to an end as the last open hearths cease production in the United States.

A key objective of the model we develop here is to explain why, even when all the characteristics of a new superior technology are known, some firms choose to postpone adoption. The fundamental assumption is that each producer adopts at a time that is individually optimal (profit-maximising) for them. By specifying exactly how the adoption decision is made on the individual level, we can present a solid microeconomic explanation for the international effect as well as the other determinants of diffusion. Our model belongs to the so-called decision-theoretic body of diffusion literature in which perfect information about the technology and adoption costs is usually assumed. Thus information-spreading is not given the same role as in epidemic models. We also assume perfect information, however in doing so our intention is not to deny the importance of information. Rather, by shifting the focus away from information we bring to the fore other determinants of diffusion. We view the decision-theoretic approach as complementary to models based on information-spreading. In particular, decision-theoretic modelling offers an alternative theoretical explanation for the international effect and our model also produces testable predictions about other determinants of diffusion.

The hypothesis that there is an international stock effect does not depend on the assumption of Cournot competition. Götz (1999) develops a model of monopolistic competition in which the payoff structure is similar to Reinganum (1981b). In the Götz model firms are small so that output decisions are not strategic however the payoffs of an individual firm depend on the actions of other firms. After a firm adopts it lowers the price it charges. This affects the overall price index and the demand function faced by all other producers. The result is a negative relationship between the fraction of firms that have adopted and per-period gross profits of an individual producer; this is the same stock effect as in Reinganum (1981b). Götz argues that since the extent of use increases with time in his model this is a case of a positive stock effect. However the reason for a positive relationship between time and further diffusion is the falling cost of adoption not the extent of use, and the stock effect as commonly defined is also negative in Götz' model. His study demonstrates that using the assumption of monopolistic competition we would be able to develop a model with a similar negative international effect as using the Cournot assumption. Because the results are very likely to be similar there is no need to replicate the modelling effort.

The chapter proceeds as follows. We begin with the assumptions and notation. The optimal output decision is discussed in section 4.3 and we show how gross profits from production depend on the extent of diffusion. We then derive the arbitrage condition, which specifies the optimal adoption date. By examining the arguments of the arbitrage condition we discuss how the extent of international diffusion among other factors affects the optimal adoption date. The discussion in section 4.5 gives rise to the hypotheses tested in the next chapter. We also examine the implications of some alternative assumptions.

4.2 Assumptions and notation

We consider the output choice of producers first without taking into account the location of the producer. Later we will specify which elements of producer heterogeneity are in fact country-level differences. We refer to the basic oxygen furnace as BOF.

The following notation is used.

N, N_C	total number of producers (in the world, in country C)
$m(t)$, $m_C(t)$	number of producers who use the BOF
c_{0i}	unit variable cost of firm i when it uses the old technology
c_{1i}	unit variable cost of firm i when it uses the BOF
$q_{0i}(t)$	quantity produced using old technology (firm i)
$q_{1i}(t)$	quantity produced using the BOF (firm i)
$\pi_{0i}(t)$	per-period gross profit flow to firm i using old technology
$\pi_{1i}(t)$	per-period gross profit flow to firm i using the BOF
$K(t)$	cost of adopting the BOF
$e_C(t)$	exchange rate of country C (in US dollars)

We make a number of simplifying assumptions.

1. Firms are heterogeneous in terms of the unit variable cost of production. For all firms, the cost is lower with the new technology:

$$c_{1i} < c_{0i} . \tag{4.1}$$

The cost advantage of the basic oxygen furnace, $c_{oi}-c_{1i}=\Delta c_i$, is a firm-specific constant and does not vary over time. This assumption rules out technological progress in either technology.

2. The number of producers in the world, N, is fixed with no entry or exit from the market over time. Producers are divided into a number of countries and they cannot move across borders so that the number of firms within each country is fixed at N_C.

3. The world demand for crude steel is a linear function of the world price:

$$P = a - bQ \tag{4.2}$$

The world price is determined in a common currency (US dollars). The amount received by each producer in own currency is determined by the (pre-determined) exchange rate of the producers' home currency against the US dollar. Because of computational issues we focus on the case without this exchange rate effect. We also experimented with an alternative demand function (constant elasticity of supply) but qualitatively this did not provide any results that cannot be obtained using the linear demand assumption.

4. The new technology is bought in a world market and the price of capital equipment, K_t, is taken as given. K_t falls continuously over time so that eventually the cost of adoption is so low that all producers use the basic oxygen furnace. The cost of adoption for firm i in country C depends on the world price and the domestic exchange rate e_{Ct}.

5. In the deterministic model, producers have perfect foresight. The variables which change exogenously over time are the price of new technology and exchange rates,

and the price of steel changes as more producers switch to the basic oxygen furnace. Producers can fully foresee all these changes. The unit costs of producers are also public knowledge.

6. Production and adoption decisions are made by individual production units which we also refer to as "firms" or "producers". Units may be part of a bigger entity however all decisions are made so that they are optimal for the individual unit.

7. The choice of technology is an all or nothing decision and as such there is no intra-firm diffusion process to consider. The switch to the new technology is irreversible.

The main difference between our model and Reinganum (1981b) is that we introduce cost heterogeneity and exchange rate effects. The effect of cost heterogeneity is essentially that if there are two firms for whom adoption at time t is profitable but only if the other does not adopt, then the one with the higher cost differential (i.e. higher gain from adoption) adopts. We find that the effect of the exchange rates on the price of steel is complex and therefore focus on the effect that the exchange rate has on the price of technology.

4.3 Output choice

The model we have in mind is one where producers located in a number of countries engage in Cournot competition in the output market but are price-takers in the market for new technology. The first section of the model concerns the choice of optimal output. We first show how the output choice is made when the location of a producer does not matter. We then take into account that the price received depends on the exchange rate of the country in which the producer is located.

The price of steel at any given point in time[43] depends on aggregate output. This is the sum of outputs produced by m adopters and N-m producers who use the old technology:

$$P = a - b\sum_{j=1}^{m} q_{1j} - b\sum_{j=m+1}^{N} q_{0j} \tag{4.3}$$

Whenever any of the N producers changes their output, the price of steel changes. This interdependence creates an opportunity for strategic behaviour because each producer understands how its output affects other producers' profits and vice versa. Consider now a producer i who is the latest to have adopted the new technology so that m=i. The (gross) profit that i makes in each period is:

$$\pi_{1i} = \left(a - c_{1i} - bq_{1i} - b\sum_{j=1}^{i-1} q_{1j} - b\sum_{j=i+1}^{N} q_{0j} \right) q_{1i} \tag{4.4}$$

Output is optimal if it maximises gross profit, given the price of steel i.e. output of all other firms. The first-order condition is

$$\frac{\delta \pi_{1i}}{\delta q_{1i}} = a - c_{1i} - b\sum_{j=1}^{i-1} q_{1j} - b\sum_{j=i+1}^{N} q_{0j} - 2bq_{1i} = 0 \tag{4.5}$$

The solution is the so-called reaction function, optimal output as a function of the output choices of other producers:

$$q_{1i}^{R} = \frac{a - c_{1i} - b\sum_{j=1}^{i-1} q_{1j} - b\sum_{j=i+1}^{N} q_{0l}}{2b} \tag{4.6}$$

Adding up the reactions functions of all adopters we have total basic oxygen output as a best response to total open hearth output:

$$\sum_{j=1}^{m} q_{1j}^{R} = \frac{m\left(a - b\sum_{j=m+1}^{N} q_{0j}^{R} \right) - \sum_{j=1}^{m} c_{1j}}{b(m+1)} \tag{4.7}$$

[43] We leave out time subscripts in this section for ease of presentation.

Using symmetry, the reaction function of producer k who uses the old technology when there are m adopters is

$$q_{0k}^{R}=\frac{a-c_{0k}-b\sum_{j=1}^{m}q_{1j}-b\sum_{j\neq k}^{N}q_{0j}}{2b} \tag{4.8}$$

Adding up all non-adopters gives

$$\sum_{j=m+1}^{N}q_{0j}^{R}=\frac{(N-m)\left(a-b\sum_{j=1}^{m}q_{1j}^{R}\right)-\sum_{j=m+1}^{N}c_{0j}}{b(N-m+1)} \tag{4.9}$$

The market is in equilibrium when each producer's output is the optimal response to the outputs of others. The equilibrium in aggregate outputs is obtained by substitution. We have

$$\sum_{j=1}^{m}q_{1j}^{*}=\frac{m\left(a+\sum_{j=m+1}^{N}c_{0j}\right)-(N-m+1)\sum_{j=i}^{m}c_{1j}}{b(N+1)} \tag{4.10}$$

$$\sum_{j=m+1}^{N}q_{0j}^{*}=\frac{(N-m)\left(a+\sum_{j=1}^{m}c_{1j}\right)-(m+1)\sum_{j=m+1}^{N}c_{0j}}{b(N+1)} \tag{4.11}$$

The equilibrium world crude steel output is

$$Q^{*}=\frac{Na-\sum_{j=m+1}^{N}c_{0j}-\sum_{j=1}^{m}c_{1j}}{b(N+1)} \tag{4.12}$$

And the equilibrium price of steel is

$$P^{*}=\frac{a+\sum_{j=m+1}^{N}c_{0j}+\sum_{j=1}^{m}c_{1j}}{N+1} \tag{4.13}$$

As the extent of diffusion increases, world crude steel output increases as the price of output falls. When i adopts the price of steel falls by $(c_{0i}-c_{1i})/(N+1)$. The equilibrium output of adopter i is

$$q_{1i}^{*}=\frac{a-(N+1)c_{1i}+\sum_{j=1}^{i-1}c_{1j}+\sum_{j=i+1}^{N}c_{0j}}{b(N+1)} \tag{4.14}$$

and the equilibrium output of k, a user of the old technology is

$$q_{0k}^{*}=\frac{a-(N+1)c_{0j}+\sum_{j=1}^{m}c_{1j}+\sum_{j\neq k}^{N}c_{0j}}{b(N+1)} \tag{4.15}$$

Because $c_{1i}<c_{0i}$ we have that $q^*_{i1}>q^*_{0i}$, that is, users of the new technology always produce a greater output than users of the old technology.

The interdependence of producers' decisions is evident in the last two terms in the numerator in (4.14) and (4.15), Σc_1 and Σc_0, which sum up the unit costs of all producers except i and k respectively. As the extent of use increases in the world, this sum falls and thus the unit's optimal output falls. In fact *in each period* except the very period in which it adopts, the producer's output falls.[44] The amount of the reduction (say, $q^*_{1i}(m)–q_{1i}^*(m+1)$) depends on the marginal adopter's cost differential (say, c_{0j}-c_{1j}) and the parameters b and N. This means that although producers are heterogeneous in their own costs, the optimal output response to a marginal adoption is the same for all producers.

Although a, b and N are parameters and therefore assumed to be fixed, it is interesting to note that an exogenous change in the parameter a (a shift of the demand curve) affects all producers in the same way – optimal output increases by the amount $\Delta a/b(N+1)$ – however a change in the number of producers N or in the

[44] A period is determined by the adoption dates of two producer who adopt consecutively. This is made explicit in the next section.

elasticity of demand b affects producers differently, and this effect depends on the producer's own unit costs.

We have shown that the representative producer's optimal response to an increase in diffusion is to reduce the output it produces. Consider now the effect on per-period gross profit. If i is the latest producer to have adopted its profit is

$$\begin{aligned}\pi_{1i} &= \left(P^* - c_{1i}\right) q_{1i}^* \\ &= \left(\frac{a - Nc_{1i} + \sum_{j=1}^{i-1} c_{1j} + \sum_{j=i+1}^{N} c_{0j}}{N+1}\right)^2 \frac{1}{b}\end{aligned} \tag{4.16}$$

And just before i adopts (i.e. m=i-1), its profit is

$$\pi_{0i} = \left(\frac{a - Nc_{0i} + \sum_{j=1}^{i-1} c_{1j} + \sum_{j=i+1}^{N} c_{0j}}{N+1}\right)^2 \frac{1}{b} \tag{4.17}$$

Profit of all producers fall with each additional adoption. This negative effect of the extent of use on profits is due to the effect of diffusion on output price. Here this effect is represented by (a reduction in) the sum $\Sigma c_1 + \Sigma c_0$. The higher a firm's own unit cost the bigger is the reduction in profit when another firm adopts. This implies that as diffusion proceeds the differences in profits across firms decrease, i.e. profits converge. Profit also falls more the greater is the cost advantage of basic oxygen furnaces (Δc) for the marginal adopter.

Although producer i's per-period profit flow falls with diffusion, its profit increases in the period in which it adopts. This is due to the fall in unit costs from c_{0i} to c_{1i} and an increase in optimal output from $q^*_{0i}(m=i-1)$ to $q^*_{1i}(m=i)$. The benefit of using the new technology today is the difference between gross profits, given by

$$
\begin{aligned}
\Delta\pi_i &= \pi_{1i}(m=i) - \pi_{0i}(m=i-1) \\
&= \frac{N\Delta c_i\left[2\left(a+\sum_{j=1}^{i-1}c_{1j}+\sum_{j=i+1}^{N}c_{0j}\right)-N(c_{0i}+c_{1i})\right]}{b(N+1)^2}
\end{aligned}
\tag{4.18}
$$

We call this the profit differential. It has a constant part which depends on i's unit costs, the number of producers N, and the demand parameters. Any demand shocks will affect the profit differential; for example, a shift-type increase (parameter a) increases the profit differential. The profit differential also changes endogenously so that with each adoption producer i's profit differential falls (as $\Sigma c_1 + \Sigma c_0$ falls). The reduction is greater the bigger is the cost differential of the marginal adopter. The fall in the profit differential also depends on i's own cost differential so that the marginal adoption has the biggest impact if i's cost differential is large. If producers adopt in the order determined by the size of their cost differential (see below), the result is quite logical because it means that the producers closest to adoption themselves have the most to lose if someone else adopts first.

There are few empirical studies of the relationship between profits and the extent of use however one such study is Stoneman and Kwon (1996). They use data on four technologies used by manufacturing firms in the United Kingdom in 1983-6, and find evidence that profits of firms using a particular technology fall as the extent of use by others increases.[45] This is consistent with the relationship suggested by the model here.

In the analysis so far we have ignored the role played by exchange rates in the output choice. We now show that the very simple model presented above becomes

[45] The empirical relationship is that the profits of users are negatively related to the number of other users and to the increase in the number of users since the firm itself adopted.

cumbersome to analyse if we take into account that the price a producer receives for output is not the world price but the price in domestic currency. Consider a representative adopter i in country a with other producers located in z different countries with different exchange rates. The price that i receives is

$$e_a P = a - b\sum_{C=a}^{z}\sum_{j=1}^{m_C} q_{1Cj} - b\sum_{C=a}^{z}\sum_{j=m_C+1}^{N_C} q_{0Cj} \tag{4.19}$$

i's reaction function is

$$q_{1i}^R = \frac{e_a\left(a - b\sum_{j=1}^{i-1} q_{1aj} - b\sum_{C=b}^{z}\sum_{j=1}^{m_C} q_{1Cj} - b\sum_{C=a}^{z}\sum_{j=m_C+1}^{N_C} q_{0Cj}\right) - c_{1i}}{2e_a b} \tag{4.20}$$

The reaction function is equivalent to (4.6) except that i's unit cost is weighted by the domestic exchange rate: $c_{1i}/2e_a b$. The aggregate reaction function for all adopters in country a is

$$\sum_{j=1}^{m_a} q_{1aj}^R = \frac{m_a\left(a - b\sum_{C=b}^{z}\sum_{j=1}^{m_C} q_{1Cj}^R - b\sum_{C=a}^{z}\sum_{j=m_C+1}^{N_C} q_{0Cj}^R\right) - \left(\sum_{j=1}^{m_a} c_{1aj}\right) \div e_a}{b(m_a + 1)} \tag{4.21}$$

Using reaction functions of this type one can obtain the equilibrium output and profits as above. The equilibrium quantities are functions of all the unit cost terms weighted by the respective exchange rate: $\Sigma c_{Ha}/e_a$, $\Sigma c_{Hb}/e_b$ etc. This implies that each unit cost term in the profit differential (4.18) is weighted by the relevant exchange rate.

The computations are straightforward but cumbersome and we find that the comparative statics results regarding the effect of the domestic exchange rate on optimal output are inconclusive even in a two-country model. This is because the domestic exchange rate has first- and second-order effects on the profit differential. The immediate effect is that the weaker the domestic currency the higher the domestic price received for output. This means that the profit differential is large, ceteris paribus. The second-order effect comes through the effect that a high profit differential has on the producer's own optimal output and consequently on the output and adoption timing decisions of all other producers, including domestically. It is this latter effect which makes it difficult to determine how exchange rates influence the profit differential.

What is clear however is that other countries' exchange rates determine how a domestic producer's profit differential falls with each additional adoption. The stronger the currency of the country in which the marginal adopter is located the greater the reduction in the profit differential (for everyone else). That is, intra-country diffusion in a country with a strong exchange rate has a greater effect on the price of steel, and thus on output and the profit differential, than intra-country diffusion in a country with a weak exchange rate.

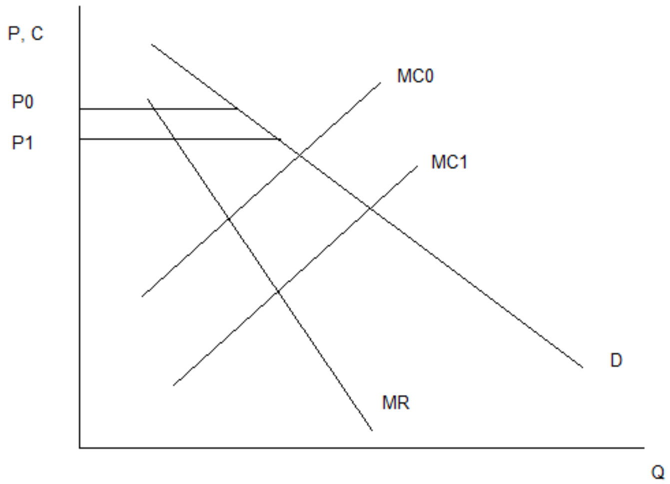
4.4 The adoption criterion / Arbitrage condition

We have shown that diffusion reduces the optimal output level for all producers and that it also reduces the benefit of using the technology. This result is the basis for the argument about the international stock effect. In this section we specify the role of the profit differential in the adoption timing decision. Because adoption is costly and the price of technology falls over time, the profit differential is only one of the factors to be considered. First we shortly discuss the assumption that the price of new technology is exogenous to diffusion.

As stated in the introduction, we assume that the price of new technology is exogenously determined. This means that the price does not depend on the number of firms who want to buy the technology; although firms behave strategically as suppliers in the output market, they do not do so as the purchasers of new technology. By adoption cost we mean, for simplicity, the price of capital equipment i.e. of basic oxygen furnaces. Adoption cost is heterogeneous only to the extent that exchange rates vary so that the domestic price paid for furnaces is determined by the world price and the domestic exchange rate.

The simplest setting in which we can talk of an exogenous price is the following: new furnaces are traded in a world market where buyers take the price as given (in US dollars). The cost of producing furnaces falls over time because of technological progress in producing furnaces. As a result, the price of furnaces falls. Exactly how price reacts to falling cost depends on the nature of competition between suppliers of furnaces. Figure 4.1 illustrates the case with a single supplier. The monopolist sets the price of furnaces so that marginal cost equals marginal revenue. As production costs fall, represented by a shift in the marginal cost curve, the monopolist reduces the price. Whether there are one or more producers of furnaces is unimportant to us; what matters is that the price of furnaces falls over time and that this price is taken as given by the producers of steel.

Figure 4.1 Price of new technology (monopoly)



Notes: P0 is the price of furnaces given cost of production MC0. As costs fall to MC1 the price of furnaces falls to P1.

Our discussion of the optimal adoption date follows Reinganum (1981b) except that we specify that the domestic price of furnaces depends on the domestic exchange rate. The representative producer i's lifetime flow of profits consists of gross profits before and after adoption and the adoption cost. The per-period gross profits are determined through Cournot competition as explained above. Producer i knows its place in the order of adopters. In the original Reinganum model producers are homogeneous and so the order is arbitrary; in our model with producer cost heterogeneity the order is determined by the firm's unit costs, as we will show. Because perfect information is assumed, producer i can work out the optimal adoption dates of all producers (including its own). Consider now i's net profits over its lifetime. The lifetime flow depends on gross per-period profits, the adoption cost paid at the time of adoption τ_i, and the discount rate r_a which we assume equals the real interest rate in country a. Lifetime net profits are expressed as the value function

$$V_{ai}(\tau_1,...\tau_i,...\tau_N)=\sum_{m=0}^{i-1}\int_{t=\tau_m}^{\tau_{m+1}}\pi_0(m)\exp^{-r_a t}\,dt+\sum_{m=i}^{N}\int_{t=\tau_m}^{\tau_{m+1}}\pi_1(m)\exp^{-r_a t}\,dt-\exp^{-r_a\tau_i}\,e_aK(\tau_i) \quad (4.22)$$

Here $\pi_0(m)$ and $\pi_1(m)$ are i's per-period profit flows which depend on the number of users m as specified above. $e_aK(\tau_i)$ is the price of new technology at time τ_i in domestic currency. Each term is discounted so that it can be evaluated in time zero. The value function can be written as

$$\begin{aligned}V_{ai}(\tau_1,...\tau_i,...\tau_N)=&\sum_{m=0}^{i-2}\pi_0(m)\int_{t=\tau_m}^{\tau_{m+1}}\exp^{-r_a t}\,dt+\pi_0(i-1)\int_{t=\tau_{i-1}}^{\tau_i}\exp^{-r_a t}\,dt+\pi_1(i)\int_{t=\tau_i}^{\tau_{i+1}}\exp^{-r_a t}\,dt\\&+\sum_{m=i+1}^{N}\pi_1(m)\int_{t=\tau_m}^{\tau_{m+1}}\exp^{-r_a t}dt-\exp^{-r_a\tau_i}\,e_aK(\tau_i)\end{aligned} \quad (4.23)$$

which makes clear that the most of the gross profit flows do not depend on τ_i but rather on other producers' adoption dates. The optimal date of adoption maximises the value function. Differentiating with respect to τ_i gives the first-order condition

$$\begin{aligned}\frac{dV_{ai}(\tau_1,...\tau_i,...\tau_N)}{d\tau_i}&=\exp^{-r_a\tau_i}\left[\pi_0(i-1)-\pi_1(i)+e_a\left(r_aK(\tau_i)-K'(\tau_i)\right)-e_a'K(\tau_i)\right]\\&=0\end{aligned} \quad (4.24)$$

where $K'(\tau_i)$ is the first derivative of the acquisition price with respect to time. Rearranging gives

$$\pi_1(i)-\pi_0(i-1)=e_a\left[r_aK(\tau_i)-K'(\tau_i)\right]-e_a'K(\tau_i) \quad (4.25)$$

which is commonly called the arbitrage condition as it refers to the costs and benefits of postponing adoption. The left-hand side is the benefit of using the new technology which we call the profit differential; we derived this quantity in the case of linear demand in (4.18). We showed that the profit differential falls with each adoption by another producer and for this reason the left-hand side of (4.25) can be interpreted as the cost of postponing adoption. On the right-hand side is the benefit of deferring adoption which is a function of the opportunity cost of adoption $rK(\tau_i)$, the rate at which furnace price falls $K'(\tau_i)$;[46] the exchange rate and any expected changes therein. A rational producer chooses τ_i so that condition (4.25) holds at the time of adoption. Note that, ceteris paribus, reductions in the profit differential discourage adoption while reductions in the benefit of postponing adoption encourage adoption.

4.5 Analysis: effects of different cost differentials and different adoption costs

We have shown how each producer chooses the amount of steel to produce and the condition which determines when to switch to the new technology. Of particular interest to us is the stock effect or the negative relationship between the optimal adoption date and adoption by other producers. Although the model is formulated on the microeconomic level, the arbitrage condition also tells us about the shape of the diffusion path in a particular country because country diffusion patterns are simply the aggregation of the choices of producers located in those countries. This is the strength of a decision-theoretic model: the model of the individual producer's choices also tells us about the determinants of intra- and inter-country diffusion. Recall that intra-country diffusion is the time path of diffusion within a country, while inter-country diffusion is the spread of a technology into new countries. The objective of this section is to develop empirically testable hypotheses about how each component of the model contributes to intra-country diffusion in particular but we also gain some insight into the determinants of inter-country diffusion.

---

[46] In reality this will refer to the fall in the price expected by firm i and may differ from the actual fall depending upon the expectations formation process.

At a high level of generality, diffusion over time can be attributed to three main factors: producer heterogeneity in production costs, the falling price of new technology, and the stock effect. The exchange rate introduces additional cross-country variation. We begin the analysis by deriving the optimal adoption date for the representative producer and examining what this tells us about the determinants of diffusion. We then examine one empirical measure of diffusion, the proportion of output produced using the new technology, and explain the difficulties in using this as a basis for empirical analysis in the next chapter. Because of these difficulties we then focus on the marginal adopter as a way of exploring the implications of the model for inter- and intra-country differences in diffusion patterns.

4.5.1 Determinants of adoption timing

The arbitrage condition can be used to solve for τ_i explicitly if we know the function $K(\tau_i)$. The functional form should be such that the cost of adoption falls over time. An example is[47]

$$K(t) = k^{-\lambda t} \tag{4.26}$$

where k and λ are some positive constants. The parameter k indicates the magnitude of the adoption cost so that a higher value implies a higher adoption cost in all time periods. The parameter λ measures the rate at which adoption cost falls and the higher the value the faster the cost falls over time:

$$\frac{dK}{dt} = -\lambda k^{-\lambda t} \ln k \tag{4.27}$$

Substituting this and the profit differential (4.18) into the arbitrage condition (4.25) gives

---

[47] This is the functional form that we assume in the next chapter.

$$\frac{N\Delta c_i\left[2\left(a+\sum_{j=1}^{i-1}c_{1j}+\sum_{j=i+1}^{N}c_{0j}\right)-N(c_{0i}+c_{1i})\right]}{b(N+1)^2}=k^{-\lambda\tau}\left(-e'_{a\tau}+e_{a\tau}\left(r_{a\tau}+\lambda\ln k\right)\right) \quad (4.28)$$

Taking logs and rearranging gives the optimal adoption date as a function

$$\tau_i=\frac{\ln\left(-e'_{a\tau}+e_{a\tau}(r_{a\tau}+\lambda\ln k)\right)-\ln\dfrac{N\Delta c_i\left[2\left(a+\sum_{j=1}^{i-1}c_{1j}+\sum_{j=i+1}^{N}c_{0j}\right)-N(c_{0i}+c_{1i})\right]}{b(N+1)^2}}{\lambda\ln k} \quad (4.29)$$

The optimal adoption date depends on i's unit costs; the extent of use elsewhere (the term $\Sigma c_0+\Sigma c_1$); the cost of adoption; demand parameters; the real interest rate; the exchange rate; and the number of producers. The expression for τ_i reveals the following ceteris paribus effects of each of the determinants on adoption timing.

1. Greater usage discourages further use because adoption of the new technology by others reduces the (marginal) benefit of use. This stock effect is reflected in the negative relationship between other producers' unit costs and the adoption date.

2. A small and declining cost of adoption encourages further use. The lower is the price of furnaces the lower is the benefit of postponing adoption. A high value of λ means that the adoption cost is falling rapidly which is an incentive to postpone adoption. Because both k and λ enter the denominator of (4.29), their values also affect the magnitude of the effect that other determinants have on τ_i.

3. Low unit costs and a big cost differential encourage early adoption. The optimal date is earlier the bigger is the cost advantage of basic oxygen, Δc_i. The level of unit costs $c_{0i}+c_{1i}$ has a negative effect on adoption so that a higher level delays adoption.

4. A high real interest rate r delays adoption because it increases the opportunity cost of adoption.

5. A weak exchange rate (high e) delays adoption and an expected strengthening of the domestic currency speeds up diffusion.

6. A more inelastic demand delays adoption. A high value of b implies that demand is less responsive to changes in the price of steel. Both gross profit level and the profit differential are lower, ceteris paribus. Because the cost of postponing adoption is lower, the optimal adoption date is later.

7. An exogenous increase in the demand for steel encourages diffusion. The larger is the demand parameter a, the earlier is the optimal adoption date. (Although the profits of non-adopters also increase, the profits of adopters increase more so the profit differential increases.)

8. The effect of the number of producers in the market, N, on adoption timing is unclear. There is no entry or exit in our model, i.e. N is fixed, so ambiguity regarding the effect of N on adoption timing is not a central concern.

Most interesting for us is the stock effect which produces a negative relationship between current usage and future adoption. The reason for the stock effect is simply the microeconomic effect of the marginal adoption on the optimal adoption date of producers that have not yet adopted. Each additional adoption reduces the price of output and thereby the profit differential. For the producer who has not yet adopted the profit differential is the cost of postponing adoption and so a reduction in the profit differential is an incentive to postpone adoption.

On the country-level the stock effect means that diffusion elsewhere discourages further usage at home, in other words, international diffusion has a negative effect on intra-country diffusion. Note that in this model the distinction between domestic and foreign use is arbitrary from the point of view of the individual adopter. That is, the stock effect concerns domestic and foreign use equally; the current extent of use, whether at home or abroad, has a negative effect on further adoptions. For the empirical researcher however the distinction is interesting because intra-country diffusion models in the literature so far have only considered domestic factors. Our model explicitly states that the extent of use elsewhere also matters for intra-country diffusion, if output is sold in a world market.

Note that the domestic stock effect only differs from the international stock effect if we compute the effect of exchange rates on the price of steel. Assuming that steel is traded in the world market not domestically, even in this case further diffusion at home affects the optimal adoption date through the world price of steel, and thus the sign of the stock effect is still negative. As explained above, comparative statics results regarding the effect of the domestic exchange rate in particular are inconclusive and for this reason we have only made explicit the effect of exchange rates on the domestic price of furnaces. This is essentially equivalent to assuming that the domestic exchange rate primarily affects the price of furnaces and that the effect of other countries' exchange rates on optimal output is negligible.

Other factors which affect the optimal adoption date of all producers equally are the world price of basic oxygen furnaces and the parameters N, a and b. While the stock effect holds back diffusion, the decline over time in the price of furnaces drives diffusion forward. The finding that λ (the speed with which furnace price falls) has a negative effect on τ_i may seem surprising at first. Consider however the producer who waits for adoption cost to fall to the level that satisfies the arbitrage condition. If price is not changing or is not expected to change i.e. $K'(\tau_i)$ is zero, then the arbitrage condition simplifies to a comparison of profit gain to the opportunity cost of adoption (the profitability criterion). The firm then acquires the new technology at the first date when the profit gain exceeds the opportunity cost of adoption.[48]

In the model we have assumed that demand for steel does not change in the sense that a and b are parameters rather than variables. However, if there is an unexpected change in the world demand for steel (higher a) this will increase the extent of use because the benefit of using the new technology increases. Indeed expressions derived earlier show that price, output, gross profits and the profit differential are all higher. A change in the price elasticity of demand also changes outputs and profits although not price. If a substitute for crude steel becomes available so that demand becomes more responsive to price (lower value of b) this increases output, gross profit and the profit differential, ceteris paribus. Postponing adoption is costlier and the extent of use increases. Note that under the assumption that the market for steel is global, a localized demand shock does not affect the local producers any differently than it affects producers elsewhere.[49]

The main factors which explain intra-country diffusion in this model are cost heterogeneity and the fall in adoption cost over time. In the firm-level diffusion literature a common approach has been to identify a source of heterogeneity in profitability, traditionally firm size, and use its distribution across firms to derive the aggregate diffusion curve. This is the so-called probit or rank approach. We have introduced firm-level rank effects into the Reinganum model through cost heterogeneity. Given the data we have at hand in the next chapter, we do not specify

[48] The opportunity cost is $r_{a\tau} e_{a\tau} k^{-\lambda\tau} - e'_{a\tau} k^{-\lambda\tau}$. If the exchange rate is also not expected to change the opportunity cost is simply the first term.

[49] With appropriate data we could test the world market hypothesis by examining whether the geographical location of a demand shock matters.

a distribution function for cost heterogeneity here. This implies that the order of adoption remains unspecified as in the original Reinganum model. However, clarity is not compromised because the implications of the model are easy to analyse. In particular it is clear how diffusion affects the optimal adoption date of the representative producer. The probit approach is helpful for explaining the identity of the adopter, however we are not particularly interested in this within a particular country and across countries we are most interested in the factors that are common to all producers within the country. The discussion in section 4.5.3 reflects a probit approach as much as our model allows.

A strength of our approach is that the model provides a framework within which the effect of institutional factors can be analysed. By this we mean variables that affect the adoption timing decision of all producers within a country. So far only two country-level factors have been specified: the real interest rate and the exchange rate. On a more structural level, the economic and political environment in which the firm operates is also expected to affect production and adoption costs and thus contribute towards explaining international diffusion. Relevant country factors also include common input costs such as the wage rate or the cost of electricity, or the cost of transport to the market. Such determinants are discussed in the context of the basic oxygen furnace in the next chapter.

4.5.2 Empirical studies

In the next chapter we test the hypothesis of an international stock effect using data on the diffusion of the basic oxygen furnace. For this purpose, we investigated whether a “structural” model can be derived from the theoretical model developed here such that it can be directly estimated using the data available to us. If one has producer-level data including the date of adoption then the arbitrage condition provides a basis for empirical study. The number of adopters in each country, m_C, is determined by the arbitrage condition that holds for the marginal adopter in that country. In the literature this approach is adopted in so-called hazard rate models. The hazard rate is the probability that a producer who has not yet adopted the new technologies does so at time t. The dependent variable is a 0/1 indicator of whether the producer has adopted, and this is regressed on variables that measure the arguments of the arbitrage condition.

The data that we have available is the annual steel output in each country, broken down by production method. In the world as a whole, the share of output produced using the new technology is obtained from (4.10) and (4.12):

$$\frac{\sum_{j=1}^{m} q_{1j}^{*}}{Q^{*}} = \frac{m\left(a + \sum_{l=m+1}^{N} c_{0j}\right) - (N - m + 1)\sum_{j=1}^{m} c_{1j}}{Na - \sum_{j=1}^{m} c_{1j} - \sum_{l=m+1}^{N} c_{0j}} \tag{4.30}$$

This expression gives the relationship between the number of users in the world m and the proportion of output that they produce. Deriving an estimating equation from (4.30) requires first that we use the arbitrage condition to express m in each country as a function of the other variables. To do this, we require some constraints on the structure of (cost) heterogeneity across producers. We investigated the simple case which corresponds to Reinganum's (1981b) original model, that is, producers are homogenous. We found that the comparative statics results are unsatisfactory even in this simple case, that is, it is not possible to make clear predictions about how the model parameters affect the share produced using the basic oxygen furnace. Adding even a very limited degree of heterogeneity makes the result is even less useful from an empirical point of view;[50] and allowing exchange rates to affect output poses further computational issues.

We therefore have to conclude that, unfortunately, we are unable to develop a structural model for use in the next chapter. However, in addition to the hypotheses already developed above, we can gain additional insight into the determinants of international diffusion by analysing what the model tells us about where the next adoption takes place.

4.5.3 Inter-country diffusion / the location of the marginal adopter

The marginal adopter is the producer who is indifferent between adopting now and postponing adoption. The objective of this section is to analyse how those elements of the costs of production and adoption that are common to all producers within a country affect inter-country diffusion. To identify the location of the marginal adopter we compare the arbitrage condition in a simplified setting in which all producers are located in either of two countries, labelled i and j. We examine three special cases each of which is simplified so that the implications of one particular type of heterogeneity become clear. We begin by assuming that there is cost homogeneity within each country; then explore the case in which the cost differential is strictly larger in one country; and finally assume that the adoption cost is strictly larger in one country.

[50] We reached this conclusion by examining the comparative statics when one of the unit cost terms is common to all producers (i.e. as in section 4.5.3.2).

4.5.3.1 Different costs of production

If the benefit of postponing adoption (the right-hand side of the arbitrage condition) is the same in the two countries then the location of the marginal adopter depends on the profit differentials in those countries. We first focus on country-level differences in the profit differential and so assume for simplicity that all adopters in country i have a unit cost c_{1i} and non-adopters have a unit cost c_{0i}, and in country j the respective costs are c_{1j} and c_{0j}. If there are m_i+m_j adopters in total, the profit differential for the m_i+m_j+1'th adopter if located in country i is given by

$$\begin{aligned}\Delta\pi_i &= \Delta\pi_{1i}(m_i+1, m_j) - \Delta\pi_{0i}(m_i, m_j) \\ &= \frac{\left(N_i+N_j\right)\left[(N_i+N_j)c_{1i}^2+(N_i-N_j)c_{0i}^2-2N_i c_{1i}c_{0i}\right]}{b\left(N_i+N_j+1\right)^2} \\ &+\frac{2\left(N_i+N_j\right)\Delta c_i\left(a-c_{0i}+N_j c_{0j}-m_i\Delta c_i-m_j\Delta c_j\right)}{b\left(N_i+N_j+1\right)^2}\end{aligned} \tag{4.31}$$

By symmetry, the profit differential of an adopter in country j is

$$\begin{aligned}\Delta\pi_j &= \Delta\pi_{1j}(m_i, m_j+1) - \Delta\pi_{0j}(m_i, m_j) \\ &= \frac{\left(N_i+N_j\right)\left[(N_i+N_j)c_{1j}^2+(N_j-N_i)c_{0j}^2-2N_j c_{1j}c_{0j}\right]}{b\left(N_i+N_j+1\right)^2} \\ &+\frac{2\left(N_i+N_j\right)\Delta c_j\left(a-c_{0j}+N_i c_{0i}-m_i\Delta c_i-m_j\Delta c_j\right)}{b\left(N_i+N_j+1\right)^2}\end{aligned} \tag{4.32}$$

A comparison of these profit differentials tells us in which country the cost of postponing adoption is the highest. The difference in profit differentials is given by

$$\begin{aligned}&\Delta\pi_j(m_i, m_j) - \Delta\pi_i(m_i, m_j) \\ &= \frac{(N_i+N_j)\left[(N_j-N_i)(c_{0j}-c_{0i})^2+(N_i+N_j)(c_{1j}^2-c_{1i}^2)-2c_{0j}\Delta c_j+2c_{0i}\Delta c_i\right]}{b\left(N_i+N_j+1\right)^2} \\ &+\frac{2(N_i+N_j)\left[(\Delta c_j-\Delta c_i)\left(a-m_i\Delta c_i-m_j\Delta c_j\right)-(c_{1j}-c_{1i})(N_j c_{0j}+N_i c_{0i})\right]}{b\left(N_i+N_j+1\right)^2}\end{aligned} \tag{4.33}$$

From the expression it is clear that unit costs matter in three ways: through the level of costs, the cost differential, and the size of the cost differential relative to the other country. The number of producers also matters both on its own and in relation to the other country. If j has the higher unit cost of the old technology and also a higher number of producers, the first term is positive and so the marginal adopter is likely to be in country j, ceteris paribus. Much of the expression is fixed if the elasticity of demand (denominator) does not change, but further diffusion changes the expression over time. As diffusion proceeds the marginal adopter is more likely to be found where the cost advantage of the basic oxygen process is higher. It is difficult to say much more about the outcome unless we restrict the heterogeneity of production costs.

4.5.3.2 Different cost differential

Assume now that producers in the two countries vary only in the size of the cost differential. We allow heterogeneity across countries in open hearth (old technology) costs, but restrict basic oxygen unit cost to be equal:[51]

$$\begin{aligned} c_{1i} &= c_{1j} = c_1 \\ c_{0i} &\neq c_{0j} \end{aligned} \tag{4.34}$$

The profit differential in country i is

$$\begin{aligned} \Delta\pi_i &= \Delta\pi_{1i}(m_i+1, m_j) - \Delta\pi_{0i}(m_i, m_j) \\ &= \frac{\left(N_i+N_j\right)\Delta c_i\left[2\left(a-c_{0i}-m_i\Delta c_i-m_j\Delta c_j\right)+N_i\Delta c_i+N_j\Delta c_j+N_j(c_{0j}-c_{0i})\right]}{b\left(N_i+N_j+1\right)^2} \end{aligned} \tag{4.35}$$

[51] Differences across countries in the cost of inputs that are used in different proportions by the new and old technologies may produce a situation in which the unit costs of either technology vary more that the costs of the other. We have in mind the price and availability of scrap metal which is used to a greater extent in the open hearth. See discussion of the historical case in the next chapter.

and similarly for country j. The difference in profit differentials across the two countries is

$$\Delta\pi_j(m_i,m_j)-\Delta\pi_i(m_i,m_j)$$
$$=\frac{\left(N_i+N_j\right)(c_{0j}-c_{0i})\left[2\left(a+c_1-c_{0j}-c_{0i}-m_i\Delta c_i-m_j\Delta c_j\right)+(N_j-N_i)(c_{0j}-c_{0i})\right]}{b\left(N_i+N_j+1\right)^2} \quad (4.36)$$

If the number of producers in each country is large, this is approximately equal to[52]

$$\Delta\pi_j(m_i,m_j)-\Delta\pi_i(m_i,m_j)$$
$$\approx\frac{(c_{0j}-c_{0i})\left[2\left(a+c_1-m_i\Delta c_i-m_j\Delta c_j\right)+(N_j-N_i-2)(c_{0j}-c_{0i})\right]}{b\left(N_i+N_j+2\right)} \quad (4.37)$$

The predictions of the model depend on the sign of the term $2(a+c_{0i}+c_1-c_{0j})$. This can be shown to be positive even without placing restrictions on the demand parameter a or the unit cost terms.[53]

We now ask where the first adopter can be found and then derive the conditions that determine where subsequent adoptions take place. The difference in the two countries in the gross profit increase available to the first adopter is given by

$$\Delta\pi_j-\Delta\pi_i=\frac{(c_{0j}-c_{0i})\left[2\left(a+c_1+c_{0i}-c_{0j}\right)+(N_j-N_i)(c_{0j}-c_{0i})\right]}{b\left(N_i+N_j+2\right)} \quad (4.38)$$

The first adoption takes place in the country where the old technology production cost is the highest, unless there are many more producers in the lower cost country.

---

[52] The approximation $\left(N_i+N_j+1\right)^2\approx\left(N_i+N_j\right)\left(N_i+N_j+2\right)$ is used.

[53] First, we show that a must be greater than any of the unit cost terms. Consider the extreme ends of the diffusion process. When all producers use the old technology, total supply is given by $Q(m_i=m_j=0)=\frac{(N_i+N_j)a-N_ic_{0i}-N_jc_{0j}}{b(N_i+N_j+1)}$. Supply is positive if $a>c_{0i}$ and $a>c_{0j}$. Conversely when all producers have adopted blast oxygen, supply is $Q=\frac{(N_i+N_j)(a-c_1)}{b(N_i+N_j+1)}$. This is positive if $a>c_1$. These restrictions on the size of parameter a imply that $2(a+c_{0i}+c_1-c_{0j})>0$.

If $c_{oj}>c_{oi}$ and $N_j \geq N_i$ then $\Delta\pi_j>\Delta\pi_i$. The first adopter is in j if there are at least as many producers in j as in i. If $c_{oj}>c_{oi}$ and $N_j<N_i$ then $\Delta\pi_j>\Delta\pi_i$ if $2(a+c_1)>(N_i-N_j-2)(c_{oj}-c_{oi})$. Thus even when there are more producers in i, the first adopter is still likely to be in j unless the difference in the number of producers is very large.

The second adopter is also in country j if

$$2(a+c_{0i}+2c_1-2c_{0j})>(N_i-N_j)(c_{0j}-c_{0i}) \tag{4.39}$$

The left-hand side is most likely positive. Therefore, if there are at least as many producers in the higher cost country then the second adopter will also be found there.

This analysis suggests that the differences in the cost advantage of new technology across countries play a crucial role in the pattern of diffusion. If producers are distributed evenly across countries then the result is similar to the case of homogeneous costs: producers in the country where the cost advantage of oxygen is the highest adopt first. The analysis also highlights the role played by the total number of producers in each country: if the domestic stock of producers is relatively large, this can overcome the effect of a smaller cost differential when the extent of use is low. That is, if one country is much larger in terms of the number of producers, this can be where the first adoptions take place even though the cost advantage is higher in the smaller country.

In general, given m_i adopters in i and m_j adopters in j, the next adopter is in country j if

$$m_i\Delta c_i+m_j\Delta c_j<a+c_1+c_{0i}-c_{0j}+\frac{(N_j-N_i)(c_{0j}-c_{0i})}{2} \tag{4.40}$$

Without knowing the size of a, it is impossible to predict which one of two possible outcomes will occur. If the higher-cost country has an equal or bigger stock of potential adopters, diffusion starts in that country and switches to the lower-cost country with all subsequent adoptions taking place there. This sequence of events may also occur if there are fewer potential adopters in the higher-cost country and the parameter a is relatively large. But if a is relatively small so that the right-hand side is less than zero, all adoptions including the first take place in the lower-cost country. That is, the technology only diffuses in the country where the old technology unit cost is the lowest.

Thus we have shown that if producers only vary in the production costs of the old technology, intra-country diffusion will not take place in both countries simultaneously and the pattern of international diffusion may be such that intra-country diffusion in one country is completed before there is any increase in inter-country diffusion. The key parameters that determine the outcome are the demand parameter a, the relative number of potential adopters and the unit costs. It is unlikely that all potential adopters adopt the new technology in this model. It is also possible that none of the producers in the higher-cost country adopt, however we also find that early on in the diffusion process the marginal adopter is more likely to be in the higher-cost country. A bigger stock of potential adopters also increases the likelihood of early adoption. These results arise if the only source of heterogeneity is the unit cost of the old technology across the two countries.

4.5.3.3 Different costs of adoption

Finally we consider the effect of country heterogeneity on the real interest rate and the exchange rate. For this purpose, assume that the costs of production are the same in both countries

$$\begin{aligned} c_{1i} &= c_{1j} = c_1 \\ c_{0i} &= c_{0j} = c_0 \end{aligned} \tag{4.41}$$

so that the profit differential is the same in both countries

$$\Delta\pi_j(m_i, m_j) = \Delta\pi_i(m_i, m_j) = \Delta\pi(m_i, m_j) \tag{4.42}$$

Suppose that the benefit of postponing adoption is higher in country i for all t:

$$e_j\left[r_j K - K'\right] - e_j' K < e_i\left[r_i K - K'\right] - e_i' K \tag{4.43}$$

Then, the arbitrage condition holds for the marginal adopter in country j,

$$\Delta\pi(m_i, m_j) = e_j\left[r_j K - K'\right] - e_j' K \tag{4.44}$$

but potential adopters in country i always strictly prefer to postpone adoption

$$\Delta\pi(m_i, m_j) < e_i\left[r_i K - K'\right] - e_i' K \tag{4.45}$$

The outcome is that all producers in the country with the lower domestic cost of adoption, country j, adopt before the first adoption occurs in country i.

This implies that adopters in the country with the lower real interest rate adopt earlier, ceteris paribus, because the opportunity cost of adoption is smaller. In this set-up this means that a high inflation rate encourages further diffusion by reducing the opportunity cost of adoption.

A second implication is that a strong exchange rate speeds up diffusion. However, if producers anticipate the currency to strengthen in the future this is an incentive to postpone adoption. Note that the full impact of exchange rates is not made explicit here; both countries' exchanges rates affect the profit differential (as a weight on the unit cost terms) but in our model the direction of these effects is unclear.[54] If the effect of the domestic exchange rate on adoption cost dominates the effect on the profit differential, as we have assumed, then producers in the country with the stronger exchange rate (smaller e) have the smaller opportunity cost of adoption and will therefore be the first to adopt, all else equal. An expected strengthening of the exchange rate (e'<0) increases the benefit of waiting, which means that diffusion is more likely to occur in the other country first.

The final implication of this exercise is that any permanent differences in adoption costs across countries are important, as are any expected changes. This includes technology-specific adoption costs but also institutional factors or structural features of the economy that affect investment decisions more generally. This argument is further pursued in the next chapter.

We conclude that persistent or underlying differences in the costs of adoption are important determinants of inter-country diffusion if producers are (relatively) homogeneous in terms of the costs of adoption. Intra-country diffusion is also expected to be fast relative to inter-country diffusion in this case.

[54] It should be noted that in the two-country model the computational difficulties are less of an obstacle however still considerable.

4.6 Alternative assumptions

In this section we look in some detail at how alternative assumptions change the results of the model in particular how the model may be modified to accommodate an endogenous adoption cost.

The main advantage of the exogenous adoption cost assumption is that it simplifies the model greatly. As the price of furnaces does not depend on the demand for furnaces we only need to model one market, the market for output. That is, we do not engage with the supply-side diffusion literature (see Stoneman 2002 and Stoneman and Battisti 2010, forthcoming). The price of furnaces can be endogenous to the diffusion process for example because of learning-by-doing, economies of scale, or strategic behaviour. The first two reduce the cost of adoption and thereby create a type of feedback loop with a positive effect on further adoption. If there are scale economies in the production of furnaces, the unit cost decreases as quantity increases. This is a very realistic scenario especially when a technology is new; the first unit is the most expensive to build. Because firms behave strategically in the output market in our model, there is no particular reason to expect that they do not do so in the furnace market as well. The result is a complicated model that is possibly a closer representation of reality but at the cost of analytical clarity.

Two types of learning can lead to a feedback loop. First, learning can occur on the buyers' side if learning from other adopters takes place. Here the argument is that late adopters have low adoption costs because they can obtain valuable information from more experienced users of the new technology. For example, Teece (1977) argues that adoption of a technology requires the transfer of unembodied technological knowledge such as engineering drawings, and as the technology diffuses this information becomes more generally available. The learning-by-doing argument and the epidemic literature have a common element in the emphasis on information. Endogenising the adoption cost therefore provides a possible way to incorporate epidemic effects into a decision-theoretic model of diffusion.

Second, learning-by-doing can take place on the supply-side. Learning spillovers mean that the cost of producing the technology (furnaces) falls with cumulative output as in David and Olsen (1986). This argument provides an interesting opportunity to endogenise the supply side of the diffusion process and thereby establish an alternative hypothesis about the mechanism by which international and domestic diffusion are linked. That is, we can argue that the direction of causality is not simply from a falling price of technology to further diffusion but that the extent of use also affects the price of technology. However, this is a step too far for the present study. Our objective here is to establish a theoretical connection between domestic and international diffusion and we propose the stock effect as this connection. Endogenising the price of technology is an alternative, and it can be also considered an extension which is an interesting direction to be pursued in the future. Furthermore, the data available to us for the empirical study would not permit to test hypotheses arising from such a model. Therefore, the endogenous adoption cost alternative is not explored further here.

Relaxing the assumptions of perfect foresight and full information is another avenue along which the model can be developed. In the first instance, uncertainty introduces the option value of waiting into the arbitrage condition (Stoneman and Toivanen 2006; also Dixit and Pindyck 1994). We have assumed that the representative producer knows exactly how its gross profits and the cost of adoption change over time. The adoption decision then involves comparing the costs and benefits of postponing adoption. A first step would be to allow the cost of adoption and its time path to be uncertain. Then the assumption that other producers' adoption dates are known in advance could be relaxed in which case decisions are based on expectations about future diffusion. As the true extent of use becomes known, beliefs about future diffusion are updated.

We consider the assumptions we have made to be justified given our twin objectives in this chapter, namely to give micro-economic foundations to the hypothesis that the extent of use elsewhere matters for domestic diffusion, and also to give a theoretical framework for the empirical analysis of a number of empirical determinants of intra-country diffusion in the next chapter. Note that we argued already in the introduction (section 4.1) that the negative stock effect does not depend on the assumption of Cournot competition. We have explained in detail the difficulties that arise from allowing heterogeneity in both production and adoption costs. Even a general model with certainty and exogenous adoption cost is difficult to solve. We recognise the value of generality but also that the more general the model, the more difficult it is to analyse. We use a specific model with an imposed structure which does not aspire to generality but allows detailed predictions about the determinants of diffusion that provide a basis for empirically testable relationships.

4.7 Conclusion

The analysis in this chapter is based on the individual producer's adoption timing decision and is an extension of the work by Reinganum (1981b). Each firm has two decisions to make over its lifetime: the amount of output it produces in each period, and the date at which it switches from the old production technology to the new one. The adoption decision is not strategic but output decisions are. Competition in the output market is of the Cournot type: optimal output depends on the output levels of all other producers, and the equilibrium is such that each producer's output is the optimal response to other firms' output levels. The price of new technology is taken as given, i.e. it is exogenous to diffusion. Price falls over time which drives diffusion forward. Strategic adoption decisions are the subject of the pre-emption literature or order models (Fudenberg and Tirole 1985) and we do not consider that here.

We extend the Reinganum model by introducing heterogeneity in production costs and by interpreting the model in an international context. We are not particularly interested in the order in which firms adopt within each country, but heterogeneity provides a way to analyse differences across countries in the diffusion process. By assuming linear demand and a particular form for adoption costs we derive the arbitrage condition and solve for the optimal adoption date. The arguments of the arbitrage condition are divided into market variables common to all producers namely the price of new technology, demand for output and the extent to which the new technology is used by producers; country-level determinants of the costs and benefits of adoption of which we have discussed the exchange rate and the interest rate; and producer heterogeneity. Analysing the determinants of intra-country diffusion on the basis of the arbitrage condition gives microeconomic foundations to the hypotheses about the determinants of international diffusion.

We propose that a link between international and domestic diffusion is established through the price of output in a so-called stock effect. The extent to which the technology is used is positively related to total output and negatively related to the price of output. The gross profit from production falls with diffusion as does the benefit of using the new technology, the profit differential. The stock effect means that adoption of the new technology by others makes adoption less attractive to those who have not yet adopted.

An important feature of our model is that the stock effect is domestic as well as international. By this we mean that further use at home has the same negative effect on the profit differential as further use abroad. As we have presented it, the model states that the size of the stock effect depends primarily on the cost differential of the marginal adopter so that the bigger is the cost reduction enjoyed by the marginal adopter, the bigger is the fall in the profit differential for all who have not yet adopted. In other words for an individual producer an adoption by a domestic competitor is as 'bad' as an adoption by a foreign competitor; both reduce gross profits from production in each period and the profit gain that the producer obtains when it adopts.

More insight to the country determinants of diffusion can be gained if we impose some structure on cost heterogeneity. A particularly useful assumption is that a proportion of the cost differential is common to all producers within a country. Reasons may be the costs of inputs such as the wage rate or cost of electricity, or the effects of institutions or policies that affect the firm's ability to reduce costs through adoption. We analysed a two-country model in which heterogeneity is restricted to cross-country differences only (section 4.5.3) and found that in this extreme case a likely outcome is that countries adopt one after the other: inter-country diffusion increases only once intra-country diffusion in one country has been completed. Combined with a degree of cost heterogeneity within a country, we have a model in which intra-country diffusion takes places simultaneously across countries and country-level heterogeneity contributes to cross-country differences.

Because of the interdependence of producers in Cournot competition, the arbitrage condition determines not only a particular producer's adoption timing but it also affects all other producers' optimal decisions. The interesting implication is that the country-specific cost factors not only affect diffusion within that country but through the stock effect they also affect international diffusion. This difficulty was discussed here in relation to exchange rates. Because the domestic exchange rate determines the price producers actually receive for their output, it affects both production and adoption decisions. There is then a second-order effect due to the assumptions of Cournot competition and perfect information, namely that all countries' exchange rates affect all producers' decisions because producers are interdependent through the market for output. The same first- and second-order effects are present when an element of the benefit of adoption is country- not just producer-specific. To illustrate, this would mean that a policy which changes the benefit of adoption in one country has a first-order effect on domestic diffusion but also a second-order effect on international diffusion as producers elsewhere react to the change in diffusion in one country.

We also discussed the other determinants of the optimal adoption date. A high interest rate and a low or inelastic demand delay adoption. A faster falling adoption cost and an expected strengthening of the domestic exchange rate also encourage producers to postpone adoption. Low unit costs of production, a high cost differential, and small adoption costs encourage adoption. Therefore diffusion is likely to occur faster in countries where the cost of adoption is small, the cost advantage of the new technology is large, costs of production are low, the real interest rate is low and the exchange rate is strong.

The stock of potential adopters is one of the two main determinants of diffusion in the epidemic framework and our results support the importance of this factor. The absolute number of producers is a parameter in the arbitrage condition and as such simply a constant term. We examined where the next adoption is likely to take place in a two-country model and found that the distribution of producers across countries matters as a mediator of the 'advantage' of a large cost differential. Generally, the marginal adopter is more likely to be found in the country where the cost advantage of oxygen is high; however, this result may be reversed if the stock of producers is considerably higher in the other country. The number of producers does not have a central role as in epidemic models or in pre-emption models in which it is the adoption timing decision which is strategic. Our finding regarding the distribution of producers does not have an obvious theoretical reason other than it is an outcome of the assumption that there is strategic interaction of the Cournot type in the output market.

The depth of the analysis which we have engaged in using a rather simple model of diffusion illustrates the power of the decision-theoretic approach more generally. Aggregate diffusion patterns can be analysed as the outcome of each producer's decisions on optimal output and adoption timing. We have chosen not to specify the distribution of heterogeneity within countries as would be done in the so-called probit approach (Davies 1979). The reason is essentially that we apply the model to a particular empirical setting in the next chapter and only country-level data is available to us for estimation. However we have demonstrated how the arguments of the arbitrage condition can be used to derive hypotheses about the country-level determinants of diffusion.

5 Empirical analysis of the diffusion of the basic oxygen furnace

Equation Chapter 5 Section 1

5.1 Introduction

In this chapter we test the key hypothesis that arises from the theoretical model developed in the previous chapter: that international diffusion affects domestic diffusion through a negative stock effect. Empirical evidence of the stock effect which more generally is defined as a negative relationship between current and future use is very scarce in the diffusion literature at large (see discussion in section 1.3). This chapter contributes to this wider discussion about the stock effect and also makes a particular argument about the empirical relationship between intra-country diffusion processes across countries.

The empirical context is the diffusion of the basic oxygen furnace (BOF), a method of producing crude steel. The BOF or basic oxygen process is, together with electric mini-mills, the main method of crude steel production today. The first successful test of the BOF took place in Linz, Austria, in 1949 and commercial production began in 1952. The idea of using pure oxygen instead of air in steel-making was not new; in fact, attempts to use oxygen to speed up the conversion of pig iron and scrap in the open hearth furnace, the dominant process of steel-making before the BOF, had taken place in a number of countries. Henry Bessemer had already recognised the benefits of using oxygen in his patent application for the Bessemer process, the predecessor to the open hearth (Maddala and Knight 1967). Early experiments using oxygen had suffered from furnaces cracking when oxygen was inserted (Adams and Dirlam 1966). Engineers at Linz were successful because oxygen was inserted through a lance from above the furnace, which prevented the furnaces from cracking.

The case of BOF diffusion is very well suited to analysis using the decision-theoretic model developed in the previous chapter. Key assumptions of the theoretical model can be considered reasonable representations of the facts. The first assumption is that the decision to use the BOF is a decision about when to replace existing technology. By the existing technology we mean primarily the open hearth furnace and also the Bessemer process where it was still used at the time. The assumption is consistent with previous academic studies (see below). Contemporary commentary from the 1969 issue of the annual OECD publication "The Iron and Steel Industry in 1968" also takes a similar view as it concludes that increases in the proportion of crude steel produced using the BOF in 1968 implied in practically all cases a displacement of basic Bessemer and open hearth steel.

A second key assumption is that the BOF is the superior technology with which the costs of production are the lowest. We discuss the cost advantage of the BOF in detail in section 5.3.3.1. Superiority implies that the users of BOF produce more than the users of the open hearth and that the gross profit of production is the highest with the BOF. This is supported by for example a summary of adoption decisions in 1966 in another edition of the aforementioned OECD publication: "Investment in most of the member countries is aimed in the main at modernization and rationalization, the replacement of obsolete plant by new units, the reduction of production costs, and the improvement of companies' competitive positions. [The] projects... can often lead to marked increases in potential supplies to the market."(OECD 1966: 74). Adams and Dirlam (1966:184 footnote) quote an article in the Wall Street Journal from the same period which states that BOF is installed to reduce costs not expand capacity.

These comments support two key features of our model: diffusion concerns the replacement of older technologies by BOF; and the adoption decision is motivated by profit.

An important assumption is also that there is no uncertainty regarding the characteristics of the BOF. From the late 1950s onwards, contemporary evidence suggests that the characteristics of the technology are widely known. There was a "great mass" of technical industry publications, and site visits to the original Austrian plants were made, so that overall producers had very good access to technical information (Adams and Dirlam, 1966:176 and references therein).

There was also plenty of data on the use of BOF around the world. The annual OECD publication "The Iron and Steel Industry" is our main data source in which detailed information on the steel industry in each of the member countries of the European Coal and Steel Community (ECSC) and the United States is presented. Other countries added in later years include the Soviet Union, Japan, Canada, other European countries, Australia, and New Zealand. Data on annual crude steel output is reported by production method. Detailed figures are also given for a large variety of steel products. The data is highly comparable across countries, consistent and reliable with very few corrections made later on. For each ECSC country, the accompanying analysis includes estimates of future demand, a discussion of investment plans and any encountered problems.

Overall the evidence suggests that contemporary comparison of 'progress' across countries was meticulous. We therefore consider it very reasonable to assume that producers had the kind of information that the theoretical model requires, including a level of detail about output which enabled producers to predict the extent to which their competitors would be using the new technology in the near future.

The assumption of a world market for steel is not intended as a literary definition of a market but what our model requires is that profits from production depend on world output of steel. We consider the international focus of contemporary discussion and contemporary academic studies to be an indication that steel producers were at least to some extent influenced by what was happening to production elsewhere. The early studies were typically concerned about the perceived decline in the competitiveness of the United States steel industry, and highlighted the relatively late adoption of the BOF by the largest producers as a central issue (see McAdams 1967, Adams and Dirlam 1966; see also Tarr 1985, Oster 1982).[55] It was suggested that oligopolistic domestic market structure may explain late adoption by the largest firms, however foreign competitive pressures were considered to be considerable. McAdams (1967:472) concludes that "United States steelmakers were influenced by foreign competition in their decision to introduce new oxygen steel capacity". Within Europe, Kipping et al. (2001:85) argue that the establishment of the European Coal and Steel Community in 1953 increased the degree of competition within Europe which "offered considerable opportunities" to companies who had invested in new technology. We consider that there is sufficient indication in the literature to suggest that the steel market was at least to some extent international in scope.

---

[55] In the wider literature, BOF represents an example of diffusion which is contradictory to the Schumpeterian hypothesis that firm size is positively related to diffusion speed (Davies 1979). Oster (1982) shows this by estimating production functions using firm-level data in the United States and she finds that large firms tended to adopt the BOF more slowly than small firms.

None of the studies of BOF diffusion so far has examined the empirical evidence in light of a decision-theoretic model of diffusion, nor studied international diffusion as we do here. We analyse a sample of 15 OECD countries over the period 1952-1985. Unlike previous studies we do not use firm-level data but aggregate crude steel production broken down by method in each of the 15 countries. The study period begins in the year the BOF is first used in Austria and ends when all other countries have stopped using the open hearth except the United States, in which the old technology was in use until 1992. The data is from HCCTAD for which the data source is various issues of the annual OECD publication "The Iron and Steel Industry".

The empirical model is based on the model presented in the previous chapter in which the arbitrage condition determines the optimal date of adoption for each individual producer. The aggregate diffusion data is then taken as the sum of producer decisions about how much steel to produce in each period and when to replace open hearths with basic oxygen furnaces. In addition to the international effect, we test a number of hypotheses concerning other determinants of diffusion based on country-level heterogeneity in costs and benefits of (postponing) adoption and determinants that are common to all producers regardless of their locations, such as the level of demand and the price of furnaces. The data available to us limits the extent to which we can test these hypotheses but our approach is to justify each explanatory variable in terms of the arguments of the arbitrage condition. The set of regressors is then based on the theoretical model, previous studies of BOF diffusion, and also some of the arguments in economic growth literature.

The chapter contributes to the literature in three ways. Firstly, we use a decision-theoretic model as the basis for an empirical study of diffusion using country-level data. Second, we seek evidence of a stock effect, in particular the proposed negative relationship between extent of use elsewhere and domestic diffusion. Third, we investigate the effect of economic and political variables not usually considered in diffusion studies such as real GDP, educational attainment and political stability. We find that there is robust evidence for the international effect but the evidence in support of the other determinants of diffusion is more suggestive. We also find that the effect of past domestic diffusion on further diffusion is positive which is contradictory to the hypothesis of a stock effect. This raises questions about how the theoretical model can be refined which are discussed in the conclusion (section 5.7).

The chapter proceeds as follows. First we present a short review of the literature regarding the costs of production using the BOF versus the open hearth, and the costs of replacing open hearth furnaces with basic oxygen ones. The empirical model is then presented and we explain how it relates to the Cournot model of the previous chapter. We present the four measures of diffusion that are used as dependent variables, discuss the estimation method (GMM) and then make a case for each of the explanatory variables. In section 5.4 we present plots of the data and discuss estimation details such as the study period and the “capacity” measure which we derive from the output data. Results are presented and discussed in section 5.5 and robustness checks in 5.6.

5.2 Cost advantage of the basic oxygen furnace and adoption costs

We begin with a review of the literature regarding the diffusion of the basic oxygen furnace. The year 1960 stands out as a turning point and for this reason we point out that, as is explained later, although we have data from 1952 onwards the actual period used in estimation begins in 1960. The literature indicates potential sources of heterogeneity across firms in the cost advantage and installation costs of the BOF. The two are related and as such difficult to discuss separately; we begin with adoption cost.

The capital costs of a new BOF were lower than those of a new open hearth furnace. According to one contemporary estimate from 1960 the installation of a complete modern open hearth shop cost about twice as much as the basic oxygen process (exclusive of the oxygen plant) with the same capacity (Hogan 1971:1523). The comparison is not straightforward because of differences in the two technologies. For example, the BOF requires a hot metal charge whereas the open hearth furnace can take a cold charge. Hogan (1971) points out that this means a BOF plant must have a blast furnace[56] supported by a coke oven nearby. Although typically blast furnaces and coke ovens were switched to support the BOF instead of open hearths, if not available nearby the capital cost of a blast furnace and coke oven would increase the costs of adopting BOF considerably.

In both contemporary and academic literature, the changing cost of adoption is discussed with reference to technological change in BOF; we have not come across

[56] The blast furnace is a method of reducing iron ore: iron ore, coke and limestone are made into pig iron (Hogan 1971: 1471).

any study of how the price of furnaces changed. It seems appropriate to think of the price of furnaces as a quality-adjusted price. Adams and Dirlam (1966:181) identify expected technological progress as the main benefit of postponing adoption in 1960. McAdams (1967) argues that by 1962 adoption costs had reduced considerably because of 'perfecting of the process' and increasing converter size, range of outputs, and maximum proportion of scrap. One problem with early BOF converters was that the linings of the converters deteriorated and required replacing very quickly, so that in practice every plant required at least two vessels which could not be operated simultaneously for any length of time (Rosegger 1980:126). There were also concerns about the suitability of BOF to producing the full range of steel qualities namely high carbon, silicon and alloy steels. By the early 1960s at the latest this uncertainty had disappeared (Meyer and Herregat 1974:151, Hogan 1971:1522). Analysts agree also that the price and availability of oxygen did not hinder the use of oxygen in steel-making. Mass production of high-quality oxygen was possible from 1929, and the price had fallen to a sufficiently low level by 1940 the latest (Adams and Dirlam 1966, Maddala and Knight 1967, Hogan 1971:1545).

The profitability of adopting BOF can be considered to have been well established by around 1960. Sumrall (1982:428 footnote) argues that by 1960 there was no doubt of the cost savings of the BOF although "prior to 1960 the evidence favoring the BOF (in a BOF versus open hearth choice) could not be taken 'without qualification'". Meyer and Herregat (1974) argue that by 1961 or 1962 all countries and firms were facing a homogenous technology and that all 'technological problems of adoption' had been resolved. Also Tarr (1985) argues that already after 1956 no new open hearths should have been built.[57]

Two factors are important as determinants of the cost advantage of BOF, i.e. the extent to which the operating costs of BOF were lower than those of the other available technologies: the heat or tap-to-tap time (or tons produced per hour), and the proportion of scrap, an input. The heat time is the time between charging the furnace with scrap and pig iron and pouring out the finished steel and this was less than 1 hour for BOF compared to 8 - 10 hours for a modern open hearth (Hogan 1971:1543, Meyer and Herregat 1974, and Rosegger 1980:121).

The faster heat time reduced unit costs. However, plants usually had to be redesigned to accommodate the faster heat time of the BOF and this increased adoption costs. McAdams (1967) argues that because of this redesign cost, adoption costs were lower for plants which currently used basic Bessemer (Thomas) furnaces than those which only used open hearths. Basic Bessemer furnaces were designed for a tap-to-tap time close to that of the BOF (about one hour). The BOF could then be adopted by simply replacing Bessemer furnaces or even using the two side-by-side. In contrast, plants which only used open hearths with a large batch size and long tap-to-tap time had to be completely revised including ladles, cranes, transportation and handling equipment in order to adopt the BOF. McAdams (1967) argues that in this case it may have been attractive to increase output by installing oxygen lances to existing open hearths instead of adopting the BOF.

[57] This is based on Maddala and Knight's (1967) argument that although the BOF could not be said to be commercially proven before 1959, it had been used commercially to such an extent that by 1956 it should have been considered in investment decisions.

Indeed doubts about the cost advantage of BOF have been voiced given the advances in open hearth technology. Hogan (1971:1523) argues that during the 1960s the use of oxygen (instead of air) in the open hearth reduced the heat time by half, so that at the time of writing the operating costs of a modern open hearth shop using oxygen were "not too far above" those of a BOF shop. This is an example of technological progress in the old technology similar to the use of steel structures in sailing ships, discussed in Chapter 3. It suggests that the reality of diffusion is more complex than what is assumed in most theoretical models of diffusion.

The cost advantage of BOF also depends on the scrap-to-metal ratio. The open hearth process is very flexible since the proportion of scrap can vary from 20 to 80 per cent. The electric arc can use charges of up to 100 per cent scrap. In contrast the BOF can take only up to 30 per cent scrap (Meyer and Herregat 1974:151). This fact has been used to argue that scrap prices may have affected the diffusion of the BOF. The key point is that electric furnaces were not good substitutes for open hearths at least not until the late-1970s (Rosegger 1980), although they can be considered the best technology for the production of highly specialised steels (Meyer and Herregat 1974). Our focus is on bulk steel and so the electric arc is not considered a substitute to the BOF. This is in line with most of the literature. Maddala and Knight (1967) argue that low scrap prices may have played a part in the Soviet Union. Beeson and Giarratani (1998) find that local electricity cost was a significant determinant of plant closures and capacity reductions in the United States in the 1970s and 1980s but electric arc usage was not. We do not consider scrap prices or electric arc usage in our model because from the literature these appear to represent a country fixed effect at most, and such fixed effects are differenced away in the estimating equation.

Estimates have been made about the size of the cost differential between BOF and the open hearth by several authors. Adams and Dirlam (1966) and Maddala and Knight (1967) argued that the cost advantage of the BOF was so large that firms should not have built open hearths, or indeed should have directly replaced them by BOFs very early on. Over a long period 1945-1970, Oster (1982) estimates production functions for four large producers in the United States and obtains a figure of 10% cost savings compared to the open hearth. Using data for a larger sample of firms she also calculates that plant cost savings from installing the BOF ranged from $3.4 to $5.6 per ton depending on the size of existing open hearths and input type. The size of open hearths matters because there are considerable scale economies in open hearth production and therefore the cost advantage of the BOF is smallest for the largest open hearths. Indirect evidence of the cost advantage of the BOF is also provided by Beeson and Giarratani (1998) who find that plants with a higher proportion of basic oxygen as opposed to open hearth furnaces are less likely to reduce capacity or close down in the United States during the 1970s and 1980s. Sumrall (1982) find that the two US firms which had installed BOF capacity early (1960) were valued more highly by financial markets than firms which installed BOF later on which, he argues, is evidence of the cost advantage of BOF.

5.3 Model

5.3.1 Estimating equation

The empirical model is based on the theoretical model developed in the previous chapter, in particular, the discussion in section 4.5. The starting point is that, as explained in section 4.5.2, we cannot derive an estimating equation directly from the theoretical model (i.e. we do not have a "structural" model). Our approach is to more generally write an estimating equation on the basis that diffusion depends on the arguments of the arbitrage condition. We choose a simple linear functional form and estimate the model using several different measures of diffusion as the dependent variable.

The arbitrage condition for the i'th producer located in country C was derived in the previous chapter (see (4.28)):

$$\frac{N\Delta c_i\left[2\left(a+\sum_{j=1}^{i-1}c_{1j}+\sum_{j=i+1}^{N}c_{0j}\right)-N(c_{0i}+c_{1i})\right]}{b(N+1)^2}=e_C\left[r_C K(\tau_i)-K'(\tau_i)\right]-e_C' K(\tau_i) \qquad (5.1)$$

This tells us that the optimal adoption date (τ_i) depends on: own costs of production (c_{1i} and c_{0i}) and particularly the cost advantage of BOF ($\Delta c_i > 0$); price of furnaces (K) and domestic exchange rate (e_C); expected fall in the price of furnaces and in the exchange rate over time; real interest rate (r_C); level and price elasticity of demand (a and b); and distribution of producers (N) across countries. The arbitrage condition determines whether a producer should adopt today or postpone adoption.

In (5.1) the extent to which other producers use the technology is measured by the sum of others' unit costs ($\Sigma c_1 + \Sigma c_0$). As argued in section 4.7, this sum depends on the decisions made by all other producers whether at home or abroad and therefore it depends on the arguments of the arbitrage condition for all other producers. Given that we do not have producer-level data on costs, an alternative representation of the extent of use elsewhere is useful. We rewrite the arbitrage condition in terms of the extent of use at time t, denoted by A_t. The life-time net profit flow of a producer less the cost of adoption at time τ_i, the value function (4.22), as a function of A_t is:

$$V_{Ci}(\tau_i) = \int_{t=0}^{\tau_i} \pi_{i0}(A_t)\exp^{-r_C t}\,dt + \int_{t=\tau_i}^{\infty} \pi_{i1}(A_t)\exp^{-r_C t}\,dt - \exp^{-r_C \tau_i}\left(e_C K(\tau_i)\right). \tag{5.2}$$

The first term is the flow of profits when i uses the open hearth and the second term is the flow after it has switched to the BOF. With each marginal increase in A_t, the per period gross profit flow decreases (except at time τ_i); this is the stock effect. The arbitrage condition is

$$\Delta\pi_i\left[A(\tau_i)\right] = e_C\left[r_C K(\tau_i) - K'(\tau_i)\right] - e'_C K(\tau_i) \tag{5.3}$$

Recall that the channel through which A_t affects profits is the price of crude steel. Ideally we would test the two relationships separately: the effect of the extent of use on the price of steel; and the effect of the price of steel on extent of use in a given country. However, we do not have data on price and therefore we test directly the relationship between use elsewhere and use at home. The potential drawback is that in the absence of empirical support for the relationship we cannot tell which of the two "links" is not supported; however this problem does not arise in practice.

In addition to the stock effect there is possibly also a positive effect if the cost of adoption is endogenous. As was pointed out in section 4.6, endogenising the supply of new technology provides a theoretical alternative to the stock effect as a way to link international diffusion to domestic diffusion. In the absence of data on the price of steel, we cannot estimate such a model and this, together with the general focus in the diffusion literature on the demand-side, has lead us to develop the stock effect hypothesis. We have limited information from the literature about adoption cost (see above) but what we have does not suggest an endogenous adoption cost so our hypothesis is that the stock effect dominates and the effect of international diffusion is negative. Indeed, the empirical evidence also supports this hypothesis.

Consider now the probability of a new adoption in country C at time t. This is the probability that the arbitrage condition holds for the marginal adopter:

$$h_{iCt} = \Pr\left[\Delta\pi_{iCt} \geq e_{Ct}\left(r_{Ct}K_t - K_t'\right) - e_{Ct}'K_t\right]. \tag{5.4}$$

where $\Delta\pi_{iCt}$ is the left-hand side of (5.1) that is, the profit differential for producer i in country C at time t. We assume that countries are heterogeneous with respect to the real interest rate and exchange rate (hence subscripts C) but that the cost of furnaces K is the same for all producers at a given point in time. Producers are heterogeneous in terms of the profit differential. Let us now give some structure to this heterogeneity. We are not interested in the order of adoption within each country so we assume that some part of the profit differential is the same for all producers within the country, although there is also a part which is entirely producer-specific.[58] Let us denote the common country-level determinants of the profit differential by $\Delta\pi_{Ct}$ and firm-level heterogeneity by u_{iCt}. We have:

$$h_{iCt} = \Pr\left[\Delta\pi_{Ct} - e_{Ct}\left(r_{Ct}K_t - K_t'\right) + e_{Ct}'K_t \geq u_{iCt}\right]. \tag{5.5}$$

On the left-hand side of the inequality is the net cost of postponing adoption which is same for all producers in country C at time t. On the right is the producer-specific net benefit of postponing adoption. If the net cost exceeds the net benefit, i adopts. This means that for units with a high value of u_{iCt} it is optimal to postpone adoption and conversely early adopters have a low value of u_{iCt}. We observe the country-level variables (the left-hand side) but not the distribution of heterogeneity (u_{iCt}) across producers within a country. Producer heterogeneity explains the order in which producers adopt; this is reflected in the intra-country diffusion pattern.

[58] It is important that there is some producer heterogeneity within each country in order to avoid a model such as discussed in section 4.5.3 where a likely outcome is that intra-country diffusion in one country is completed before any further inter-country diffusion takes place.

We more generally write an estimating equation on the basis that diffusion depends on the arguments of the arbitrage condition. We do not have data on all of the arguments of (5.5) but in section 5.3.3 we explain the set of regressors. The estimating equation is the linear relationship

$$\begin{aligned} S_{Ct} &= \alpha_1 S_{Ct-1} + \alpha_2 A_{-C,t-1} + \beta_1 \ln GDP_t + \beta_2 PYR_{Ct} + \beta_3 SYR_{Ct} + \beta_4 PIN_{Ct} \\ &+ \gamma_1 r_{Ct} + \gamma_2 \Delta CPI\%_{Ct} + \gamma_3 e_{Ct} + \gamma_4 \Delta e_{Ct} + \gamma_5 K_t + \gamma_6 \Delta K_t + \gamma_7 t + \eta_C + \varepsilon_{Ct} \end{aligned} \tag{5.6}$$

The dependent variable S_{Ct} is one of the four domestic diffusion measures discussed in the next section. $A_{-C,t-1}$ is the extent of use elsewhere (i.e. in countries other than C). It is therefore determined by the arbitrage condition that holds elsewhere. Taken together, S_{Ct-1} and $A_{-C,t-1}$ measure the extent of international diffusion at time t-1. Both α_1 and α_2 are excepted to be negative because an increasing extent of use reduces the benefit of adoption (the profit differential). This is the so-called stock effect, which we have divided here into the domestic and the international stock effects. In chapter 3 section 3.3.2 we suggested that an appropriate measure of international diffusion is either the average extent of use elsewhere, or the level of use elsewhere less the domestic contribution. Here we use both measures: the proportion of basic oxygen elsewhere (denoted by A%) and the basic oxygen tonnage elsewhere (denoted by A). Note that these are country-specific variables.

Three of the regressors are common to all countries: ln GDP_t which is real world GDP which we use as a proxy for the demand for steel (specifically parameter a), K_t and ΔK_t which are a nonlinear function of time and its first difference, and t which is a linear time trend. These last three terms we include as a proxy for the price of adoption. Six more variables measure country differences in the arbitrage condition. e_{Ct} is the domestic exchange rate (in domestic currency per USD), r_{Ct} is the nominal interest rate, and $\Delta CPI\%_{Ct}$ is the percentage change in the consumer price index. Finally we use PYR_{Ct} and SYR_{Ct}, measures of human capital (years of primary and secondary schooling) and PIN_{Ct}, a measure of political instability, to indicate country-level differences in the profitability of adoption. The hypotheses and measurement issues are discussed in section 5.3.3.

The country effects η_C and the time-variant term ε_{Ct} are not observed by us. The country effects capture heterogeneity in the arbitrage condition across countries that is common to all producers within a country, constant over time and not captured by the other regressors. The term ε_{Ct} reflects two phenomena, between which we cannot distinguish: the distribution of producer heterogeneity within the country, i.e. the proportion of the profit differential that is entirely producer-specific u_{iCt}, and unexpected shocks to either the cost or the benefit of postponing adoption which affect the optimal adoption date of all producers within a country. Except for the shocks these terms are assumed to be known to all producers although unobserved by us.

We assume that all regressors are exogenous except for the two lagged variables S_{Ct-1} and $A_{-C,t-1}$. The country effects create serial correlation and to eliminate this we take first differences

$$\begin{aligned}\Delta S_{Ct} &= \alpha_1 \Delta S_{Ct-1} + \alpha_2 \Delta A_{-C,t-1} + \beta_1 \Delta \ln GDP_t + \beta_2 \Delta PYR_{Ct} + \beta_3 \Delta SYR_{Ct} + \beta_4 \Delta PIN_{Ct} \\ &+ \gamma_1 \Delta r_{Ct} + \gamma_2 \Delta\Delta CPI\%_{Ct} + \gamma_3 \Delta e_{Ct} + \gamma_4 \Delta\Delta e_{Ct} + \gamma_5 \Delta K_t + \gamma_6 \Delta\Delta K_t + \gamma_7 + \Delta \varepsilon_{Ct}\end{aligned} \tag{5.7}$$

Two of the explanatory variables, ΔS_{Ct-1} and $\Delta A_{-C,t-1}$, are correlated with the error term through ε_{Ct-1} even if ε_{Ct} is not serially correlated. For least squares estimation this means that two of the first-order conditions for minimisation of the least squares criterion drop out. Identification requires additional moment conditions which can be provided by instrumental variables.

We briefly illustrate the instrumental variables approach. Consider a simple cross-section regression with two regressors of which x_{1i} is endogenous: $y_i = x'_{1i}\beta_1 + x'_{2i}\beta_2 + \varepsilon_i$. There are two parameters to be estimated but only one first-order condition, $(1/N)\sum_{i=1}^{N}\left(y_i - x'_{1i}\beta_1 - x'_{2i}\beta_2\right)x_{2i} = 0$, because $E[\varepsilon_i \cdot x_{1i}] \neq 0$. An instrumental variable is some z_i that is correlated with x_{1i} but not with ε_i, that is, $E[\varepsilon_i \cdot z_i]=0$. Then we have an additional moment condition $(1/N)\sum_{i=1}^{N}\left(y_i - x'_{1i}\beta_1 - x'_{2i}\beta_2\right)z_i = 0$ and the model is just identified.

In (5.7) we require at least two instruments for the two endogenous variables, ΔS_{Ct-1} and $\Delta A_{-C,t-1}$. We use S_{Ct-2} and $A_{-C,t-2}$ and earlier lagged levels up to the fifth lag. S_{Ct-2} and $A_{-C,t-2}$ are valid instruments if $E[(\varepsilon_{Ct}-\varepsilon_{Ct-1})S_{Ct-2}]=0$ and $E[(\varepsilon_{Ct}-\varepsilon_{Ct-1})A_{-C,t-2}]=0$. This requires that $A_{-C,t-1}$ and S_{Ct-1} are predetermined and that the error terms in the levels equation (ε_{Ct}) are not serially correlated i.e. $E[(\varepsilon_{Ct}-\varepsilon_{Ct-1})(\varepsilon_{Ct-2}-\varepsilon_{Ct-3})]=0$. For $A_{-C,t-1}$ to be predetermined means that it is not correlated with the current or future errors, that is $E[A_{-C,t}\cdot\varepsilon_{Cs}]=0$ for s>t. We assume that this is true, as it is for S_{t-1}. The lagged value $A_{-C,t-1}$ is correlated with all past errors because of the way domestic diffusion affects the extent of use elsewhere.

Note that GMM and OLS are not consistent if we were to use the current extent of use elsewhere, $A_{-C,t}$, as a regressor in the model. $A_{-C,t}$ is simultaneously determined with S_{Ct} and therefore $E[A_{-C,t}\cdot\varepsilon_{Ct}]\neq 0$ and $A_{-C,t}$ is not predetermined. The very essence of the Cournot assumption is that $S_{C,t}$ influences $A_{-C,t}$ and the other way around.

Additional correlation between $A_{-C,t}$ and ε_{Ct} arises from any common shocks not observed by us. These affect both foreign and domestic adoption decisions and create contemporaneous correlation between $A_{-C,t}$ and ε_{Ct}. However $A_{-C,t}$ is not correlated with future errors and so the lagged value $A_{-C,t-1}$ is predetermined, $E[A_{-C,t-1} \cdot \varepsilon_{Ct}]=0$.

The method of estimation is one-step GMM using heteroskedasticity-consistent standard errors. We use one-step rather than two-step GMM because the two-step estimates have a considerably larger standard error (up to ten times the standard error of the one-step estimates).[59] Although the one-step GMM estimator is consistent the two-step estimator is theoretically more efficient which is why the latter is more common in empirical studies. However, some authors have found one-step estimates to be preferable (e.g. Judson and Owen 1999). Cochrane (1996) compares the choice between 1- and 2-step GMM to the choice between OLS and Generalized Least Squares and argues that one should make sure that the latter are not too different from the former. The two-step estimator is problematic if the covariance matrix of the sample moments, which is used to construct the second-stage weighting matrix, is poorly measured in the first stage. Although the one-step estimator is not efficient it is consistent. For this reason in a situation in which the two-step estimates are considerably different we consider the one-step estimates to be a justified choice.

[59] An implication is that measures of fit such as the equation standard error and the residual sum of squares are very large for two-step estimates compared to one-step estimates.

5.3.2 Dependent variables

Because our model is not a structural one there is no clear choice of dependent variable. We solve the problem by estimating the model using four different dependent variables (S_{Ct}) each of which is related to the arbitrage condition. A priori we give each set of results the same importance.

The first measure is a straightforward measure of the extent of use, the BOF tonnage produced each year. With this dependent variable we use total BOF tonnage elsewhere (denoted by A later), as a measure of the extent of use elsewhere. The second dependent variable is the BOF proportion, that is, the proportion of BOF of total open hearth plus BOF tonnage. A similar measure is used e.g. by Oster (1982). More generally, proportional measures are used in most studies of intra-country diffusion because the upper and lower boundaries are clearly defined and such a measure lends itself well to the fitting of diffusion curves. An advantage is also that there are no units of measurement (e.g. tons). As a measure of use elsewhere in this regression we use the proportion of BOF in total tonnage elsewhere (A%). Both BOF tonnage and BOF proportion are expected to be determined by the same set of explanatory variables through the arguments of the arbitrage condition. However the strength of the relationships, including the international effect, may well vary. The theoretical model gives little indication about such variation except that we expect BOF proportion to be less sensitive to changes in the demand for steel.

Although in the Cournot model demand for steel is assumed to be constant (represented by parameters a and b) in reality this was not the case. In the model, total tonnage increases only as a result of the diffusion process, i.e. as more producers switch to the BOF. The marginal adopter's output increases as it adopts to which other producers respond by reducing their output. During our study period production capacity changed not only due to this displacement of the open hearth by the BOF – the diffusion process we are interested in – but also due to considerable changes in demand. During the 1950s and 1960s, annual world steel production increased more than three-fold. Production peaked in 1974 and a recession followed immediately. There was a partial recovery in the late 1970s mainly attributable to developing countries, but another decline in 1982. Technological progress in industries using steel also contributed to a decline in demand (Hogan 1994:2-3 and Hogan steel archive 2008). In the Cournot model, such changes in demand are not modelled but can be represented by a change in the parameters a and b. The first-order effect of an increase in parameter a (a shift in demand) is that both open hearth and BOF output increase. The second-order effect is that the optimal adoption date is earlier, so BOF output increases more. There is also a third-order effect because other producers respond to increased extent of use by reducing output.

We take into account changes in the level of demand in three ways. First, we use BOF proportion in addition to BOF capacity as a dependent variable. This is expected to be less sensitive to changes in demand because open hearth output also responds in the same direction. Second, we do not use the output data directly but apply a smoothing process and construct a measure which we call capacity. This is explained in Section 5.4 where we also plot the output data against the smooth measure. Third, we include world GDP as a proxy for changes in demand.

In addition to the two level measures of diffusion which we have now proposed to use as a dependent variable, we want to also estimate the model using a measure of change in the extent of use. We want a measure that reflects the displacement of open hearths rather than changes in demand for steel, as far as possible. A measure that has been used by Maddala and Knight (1967), Meyer and Herregat (1974) and Tarr (1985) is the change in BOF tonnage divided by the change in total tonnage. This measure is motivated by a particular interest in whether producers behaved "optimally" and built BOFs when capacity was expanded and so it is not the right measure for our purposes. Another possible measure is the growth rate of BOF tonnage. We estimate the model using BOF growth rate as the dependent variable and find that although our main hypothesis of the negative international stock effect is supported, the model fits poorly. We suggest this is because the measure is particularly sensitive to fluctuation in demand and that the weak results possibly reflect the shortcomings of our proxy for demand, world GDP. We discuss the results shortly in section 5.6 but otherwise concentrate on the following two measures of change. In our view these better capture the essence of diffusion as the displacement of open hearths.

The third measure is the annual change in BOF tonnage as a proportion of open hearth tonnage:

$$\frac{CBOF_t - CBOF_{t-1}}{COH_{t-1}} \tag{5.8}$$

Here "CBOF" and "COH" refer to the capacity measure explained in section 5.4. We refer to (5.8) as the relative rate of growth which emphasises the role of the remaining open hearth capacity as the stock of potential adopters. The final measure is the negative growth rate of open hearth tonnage:

$$\frac{-(COH_t - COH_{t-1})}{COH_{t-1}} \tag{5.9}$$

Conceptually this measure only makes sense when open hearth output is falling and when that decline is due to the replacement of open hearths by basic oxygen furnaces. We find that the open hearth peak is reached in many countries only after BOF production has started and therefore the sample size is the smallest when (5.9) is used as the dependent variable.

We consider that measures (5.8) and (5.9) are interesting because they are conceptually closest to the so-called hazard rate. We have already encountered the hazard rate; it is the probability that the arbitrage condition holds, (5.4). The hazard rate is defined as the probability that a producer who has not yet adopted the BOF adopts it in a given small time interval {t,t+dt}:

$$h(t) = \lim_{dt \to 0} \frac{\Pr\{t \le \tau < t + dt \mid \tau \ge t\}}{dt} \tag{5.10}$$

where τ denotes the date of adoption. The hazard is a particularly useful measure when the date of adoption for each firm is known to the researcher. Then, the hazard provides a way to derive the estimating equation directly from the arbitrage condition. Denote the distribution function of adoption dates by $\Pr[\tau<t]=S(t)$ for $t\geq 0$ and the probability density function by $dS(t)/dt=s(t)$. The hazard is

$$h(t) = \frac{s(t)}{1 - S(t)} \tag{5.11}$$

The proportion of producers that have adopted, $m(t)/N(t)$, is an estimator of $S(t)$ for $t\geq 0$ and the change in the proportion, $d\,[m(t)/N(t)]\,/dt$, is an estimator of $s(t)$. If the number of producers is constant, $dN(t)/dt=0$ (as we assume), the hazard can be estimated by the change in the number of adopters divided by the stock of producers who have not yet adopted:[60]

[60] This expression is also the starting point of the seminal model by Mansfield (1961) although he does not develop it from a probabilistic argument. Mansfield uses (5.12) to derive the logistic model (3.2) by stating that the proportion of firms who adopt in the interval (t-1, t) is a function of the proportion who have already adopted (m/N), the profit differential, and the cost of adoption. This

$$h(t) = \frac{dm(t)/dt}{N - m(t)}. \qquad (5.12)$$

As stated earlier we do not have data on the number of users, only the output produced. Each marginal adopter adds one to the number of adopters m but the change in aggregate basic oxygen output is $\sum_{i=1}^{m} q_{1i} - \sum_{i=1}^{m-1} q_{1i}$. The magnitude of this change depends on the cost differential of the marginal adopter (whose output increases) and the amount by which other producers reduce their output (as an optimal response). Producer heterogeneity means that there is no direct relationship between (5.12) and the (shares of) output produced with each technology in the world.

More generally, (5.12) gives rise to the hypothesis that the probability that a producer who currently uses the open hearth switches to the BOF today is empirically related to the capacity that can be converted, i.e. the "stock" of potential adopters N-m(t), and the observed change in the extent of use, dm(t)/dt. The denominator in (5.8) and (5.9) is the amount of old technology capacity that has not yet been displaced. The numerators measure the change in use as the increase in BOF capacity in (5.8), and as the fall in open hearth capacity in (5.9). The extent to which these numerators are good measures of the extent of displacement is left to be empirically determined. We consider that because there is no theoretical reason to prefer one dependent variable over the other, it is appropriate not to determine a priori which of the four dependent variables is to be used. It turns out that our hypothesis regarding the international effect is supported by all the four sets of regressions which in our view confirms the appropriateness of our approach.

---

link demonstrates the close connection between Mansfield's model and the decision-theoretic approach to diffusion.

5.3.3 Explanatory variables

In this section we discuss the set of regressors, how these are related to the arbitrage condition, and how we measure them in practice. We begin with the cost differential and the costs of adoption, and then the proxy for demand, world GDP. The wider diffusion literature is then used to propose that human capital and political instability may bring about country-level heterogeneity in production (or adoption) costs.

5.3.3.1 Cost advantage of BOF and adoption costs

As discussed above, there is little information in the literature about the time path of adoption costs. Colombo and Mosconi (1995) argue that a variable which indicates the passage of time reflects changes in the price of adoption if that is not captured by any other variable in the model. We use this approach and include time effects as a proxy for adoption cost. The constant term γ_5 in the first-differenced equation (5.7) corresponds to a linear time trend in the level equation (5.6). There is no a priori reason to expect the price of adoption to fall linearly over time and therefore we include both a linear and a non-linear control.

We experimented with arbitrary convex functions of time, that is, functions which fall at a decreasing rate over time. A convex function is expected if the reason for a fall in price is economies of scale in producing furnaces, for example. We used the functional form $K=k^{-\lambda t}$. This was also used in section 4.5.1 where we solved for the optimal adoption date explicitly. Empirically, k=2 and λ=0.1 appears to perform well and this is the form we use, that is, $K=2^{-0.1t}$. We expect the coefficient on K to be negative to reflect the hypothesis that a high adoption cost discourages diffusion and a fall in adoption cost over time drives diffusion forward. The convexity of K implies that postponing adoption by one period becomes less attractive over time, as the rate at which adoption cost falls slows down.

In some regressions the coefficient on K is positive, which indicates that the time control is not picking up the effect that we were expecting. Therefore, we also estimate these regressions using biannual time dummies instead of K. The usual approach in panel data studies is to use annual time dummies. However, this is not feasible for us because the study period is so long that annual dummies end up explaining most of the variation in the dependent variable leaving little to be explained by other variables. We also find that annual dummies lead to singularity problems. We experimented with 4-year dummies but this was found to be too long a period, and so we use biannual time dummies.

We also consider two sources of heterogeneity in adoption cost across countries: the exchange rate e and the real interest rate r. Hypotheses are derived directly from the theoretical model.

A high interest rate delays adoption because it increases the opportunity cost of adoption. We divided the real interest rate effect to the nominal rate and the annual percentage change in the consumer price index. We do not use a single measure because during the study period negative values of the real interest rate are observed and these are difficult to interpret as intended in the theoretical model. The nominal rate is expected to have a negative coefficient and CPI change is expected to have a positive coefficient, and the two should be at least jointly statistically significant. Our data source is the International Monetary Fund's (1998) International Financial Statistics, and the data was obtained through ESDS International.

No single measure of the nominal interest rate is available for all 15 countries. We considered three alternative measures – government bond yield, central bank discount rate, and money market rate – and chose government bond yield because this has the best coverage in our sample. Data is available for all countries except Finland for at least some part of the period. Data is missing from 1952 onwards for Austria (until 1965), Germany (until 1956), Japan (until 1966), Luxembourg (until 1970), and Spain (until 1979). This lack of data reduces the sample size because countries are only included when interest rate data is available. The inflation measure, percentage change in the consumer price index, is computed as the annual change in the CPI series. This measure is available for all countries for the whole period.

While steel is exported, furnaces are imported. The domestic exchange rate has contradictory effects on optimal adoption timing through these two channels in the sense that a weak exchange rate makes furnaces relatively expensive but the gross profit of production relatively large. We argued in Chapter 4 that the overall effect is unclear because other firms take into account the effect of the domestic exchange rate on the profit differential, and so the domestic rate features in the decision-making process of foreign firms as well. In turn, their decisions affect domestic firms' decisions. This interdependence arises from the assumption of Cournot competition in the market for steel. The first-order effect of a weak exchange rate is that the profit differential is large[61] which means that the incentive to adopt now is high, ceteris paribus. However, producers in other countries anticipate this, and as a result the international extent of use today may be higher than if the domestic currency was stronger. That is, a weak domestic exchange rate has a second-order effect which is likely to have the opposite sign to the first-order effect (that is, the profit differential is smaller). The outcome is that the relationship between exchange rates and diffusion is so complicated that we are unable to obtain comparative statics results even in a two-country model.

The way we proceeded in the previous chapter is that we make the simplifying assumptions that the domestic exchange rate mainly affects the price of furnaces rather than the price of steel, and that the effect of other countries' exchange rates on domestic diffusion is negligible. In this scenario, we expect that the weaker the currency, the greater the incentives to postpone adoption. The domestic exchange rate is measured as domestic currency per US dollar[62] and so we expect the exchange rate to be negatively related to domestic diffusion. The data is from the Penn World Tables version 6.2 by Heston, Summers and Aten (2006).

5.3.3.2 Demand for steel and distribution of producers across countries

There were considerable movements in the demand for steel during our study period however our theoretical model is not particularly suited for analysing such changes because demand is represented by two parameters, a and b. We attempt to control for changes in demand by including a proxy for the parameter a, which represents

[61] The profit differential rises with the level of gross profits (see Chapter 4).

[62] The exchange rate is measured as units of domestic currency per US dollar. Using SDRs (the IMF preferred basket currency) does not change the results because during this period the SDR is equal to or very close to the US dollar.

the level of demand. An increase in a is expected to have the following effects on output and diffusion. First, the equilibrium outputs of all producers whether using the BOF or the open hearth change by a constant amount. If demand increases, all producers produce more. The second-order effect concerns diffusion: an increase in demand implies a higher profit differential and thus a higher extent of use, ceteris paribus. Finally the third-order effect is that equilibrium output levels fall as a response to higher diffusion. This last effect is likely to be small enough so that we expect both BOF and open hearth output to increase, although BOF output is expected to increase more, ceteris paribus.

As a proxy for changes in the demand for steel we use the logarithm of world real GDP, computed as the sum of real GDP in the 15 countries in our sample. The hypothesis is that an exogenous increase in demand, proxied by an increase in GDP, increases the payoff from adoption and so the optimal adoption date is earlier, ceteris paribus. We thus expect a positive coefficient on GDP. The purpose of proxies in general is to use an indicator that drives the majority of the variation in the original data (Durlauf et al. 2005); we expect that steel demand and aggregate demand in the economy are sufficiently correlated.

The total world stock of potential adopters N is part of the constant term that is common to all producers across countries. In section 4.5.3.2 we found that the distribution of producers across countries also matters. In particular, if a country has a large share of all producers – the stock of potential adopters – this can be an "advantage" in inter-country diffusion so that the technology is adopted earlier than otherwise indicated by the relative size of the cost differential in that country. Since the number of producers is assumed constant, it is part of the country fixed effect η_C that is eliminated when we take first differences and therefore not observed by us.

5.3.3.3 Human capital and political instability

Institutions such as the regulatory and educational systems are part of the "environment" of the firm (Teece et al. 1997). Differences in the environment across countries and over time can explain differences in either adoption cost or the benefits of adoption. We consider two such measures: average years of schooling and political instability.

The link between education, diffusion and growth has been studied since Nelson and Phelps (1966) proposed that education reduces the risks involved in adopting new technology. They argued that education of managers and scientists matters because the former make decisions about adoption and the latter seek information about inventions and distinguish the profitable ones from the rest. Nelson and Phelps (1966) linked diffusion to growth by arguing that the greater the level of education in the economy the smaller the gap between the stock of technological knowledge (the "theoretical" level of technology) and the actual level that determines aggregate output of the economy. Education lowers the costs and uncertainty of adoption and is thus expected to have a positive effect on diffusion.

In his review of both economic and sociological literature on diffusion, Hall (2004) argues that the cost of adoption "includes not only the price of acquisition, but more importantly the cost of the complementary investment and learning required to make use of the technology. Such investment may include training of workers and the purchase of necessary capital equipment... the need for complementary investment, especially for complex modern technology that requires the re-organization of the process that will use it". The need for complementary investment increases the cost of adoption including the time it takes to realise the benefits of adoption. In a similar vein, Teece (1977) argues that adoption of a technology requires a transfer of unembodied technological knowledge. The more generally available the technology, the lower is the technology transfer cost (e.g. because engineering drawings are generally available). Teece argues that labour skill level is expected to reduce this cost. More generally, the hypothesis that arises is that the overall costs of adoption are negatively related to the skill levels of workers and managers. In relation to the arbitrage condition this implies that the higher is the skill level the smaller is the benefit of postponing adoption and thus we expect a positive relationship between education and the extent of use across countries and time.

Abramovitz' (1986) concept of "social capability" provides a further basis for including education in the model. Abramovitz (1986) argued that a country requires some minimum level of development to be able to exploit new technologies. This has lead to education measures being used as proxy for social capital, that is, the social networks through which information is distributed, in the sociological and marketing literatures. Education is viewed as increasing the ability of "change agents" to process and understand the value of the new technology (Rogers 1995). Skinner and Staigler (2005) are influenced by this view in their study of the use of a number of technologies across countries. They find that education is highly correlated with the extent of use and they argue that this is so because education is a good predictor of social participation, political participation and other factors that are conducive to interaction and thus information exchange.

Since the late-1980s there has been an increasing interest in the modelling of human capital as a characteristic of an economy relevant to economic growth. Nelson and Phelps' (1966) study has been influential in this new growth literature however the proposed link between education and diffusion does not rely on uncertainty or lack of information. This is important since our model assumes the characteristics of the technology are public knowledge. The general argument is that education matters because it is a measure of human capital. Caselli and Coleman II (2001) distinguish between two views in the literature: the skill-bias view which asserts that only adoption of skill-biased technologies depends on human capital, and a more general view that adoption of any new technology depends on human capital. In the former case, the crucial factor is whether the new technology is a substitute for skilled labour. In our case, the consensus in the literature is that the BOF represents neutral technological change which means that there is no labour-saving bias (Maddala and Knight 1967; Oster 1982). Thus it is the more general "skills in adoption" – view which suggests that education matters in BOF adoption. Skills are a resource, an input to the production of goods and services, but in endogenous growth models the production of knowledge is also an integral part of growth. Thus in this literature various measures of education (enrolment, attainment, years of schooling), typically distinguished by the level of education (e.g. primary, secondary, tertiary), are included as regressors in growth models in an attempt to capture variation in human capital.

Previous empirical studies which have found evidence of a relationship between education and diffusion include: Caselli and Coleman II (2001), who study cross-country differences in computer-technology adoption; Kiiski and Pohjola (2002) who examine the average years of schooling in a Gompertz model of internet diffusion; and Perkins and Neumayer (2005), who find that secondary school enrolment increases the speed of intra-country diffusion (the length of time between first adoption and the "saturation" level) for one of the three technologies they study.

Our hypothesis is then that human capital, which we measure by the average number of years of schooling, is positively related to diffusion. The data is from Barro and Lee (2000) which uses data from UNESCO annual yearbooks. We use two measures, years of primary schooling and years of secondary schooling, computed for the population aged 25 and above. The measures are available for 5-year intervals from 1960 onwards. To have annual data we impute random values between the two closest data points and the imputations are ordered. The ordering is determined by the end points of each five-year interval; there is a strong increasing trend in secondary schooling in most countries but the trends are weaker in primary schooling. This approach to imputations avoids creating any artificial jumps in the time series at the five-year intervals. Primary schooling varies less than secondary schooling in this sample. The minimum number of years of primary schooling is 3.0 and the maximum is 6.8 with a standard deviation of 0.98. The minimum years of secondary education is 0.4, the maximum is 2.4 and the standard deviation is 1.18. Secondary education increases considerably for all countries during the period except in Germany where the level is relatively high throughout. The countries which start with a value near zero (less than one year of secondary schooling) are Italy, Netherlands and Spain.

Together with the educational system, political institutions are part of the environment in which adoption and production decisions are made. Government ownership and planning has been considered in some studies of BOF diffusion (Tarr 1985, Kipping et al. 2001). We consider more interesting the wider argument attributed to Huntington (1968) that maintaining political order is more important for economic development than the particular set of institutions in place. It is on this basis that we include a measure of political instability in our empirical model. Because the theoretical model assumes certainty, we do not take political instability as a measure of uncertainty about the political order or of property rights. Instead, we interpret instability as a contemporaneous effect on the one hand – as Przeworski et al. (2000: 188) put it: "Political upheavals divert resources and energies away from production and thus affect the contemporaneous growth of the economy" – and an indicator of structures such as the stability of the political regime on the other hand. The latter view draws on Acemoglu and Robinson (2001) who develop a model in which redistribution policies (marginal tax rate) depend on the likelihood of coups (by the elite) and revolutions (by the poor). They assume that a proportion of output is lost during coups and revolutions, and that profitability of investment is higher in democracies than in non-democracies. One of their results is that expectations about the durability of a democracy can be self-reinforcing through the impact that such expectations have on investment decisions: if firms expect democracy to persist (there is no coup) the expected profitability of investment is high and high investment in turn increases the durability of democracy.

If, as Acemoglu and Robinson (2001) suggest, high inequality leads to political instability, the degree of instability also reflects the degree of inequality. Alesina and Rodrik (1994) develop an endogenous growth model in which the tax rate depends on the degree of income and wealth inequality. In less equal societies, the demand for redistribution of wealth is strong and so inequality increases the probability of social unrest. A high tax rate reduces the rate of economic growth in this model which produces a negative relationship between inequality and growth, of which Alesina and Rodrik (1992) present empirical evidence.

We expect political instability to increase the incentives to postpone adoption of BOF. The data is from Barro and Lee (2000) for which the original source is Banks (1979) Cross-National Time Series Data Archive. The measure of instability that we use is an annual average number of assassinations and revolutions (both successful and unsuccessful) per 1 million population ("PINSTAB").[63] A value of 0 indicates stability. The measure is computed over 5-years periods and in order to have annual observations we use the average annual value for each year within the 5-year period. The data is available for each of the 15 countries for the period 1960-1984. PINSTAB is closely related to a measure "REVCOUP", the average number of coups and revolutions in the period 1960-1984, which has been used for example by Hoover and Perez (2004) who find evidence that REVCOUP is a significant determinant of growth.

5.4 Data

In this section we present plots of the four dependent variables and the two measures of international diffusion. We also explain how we have constructed a measure of "capacity" from the output data, and how data for four countries was

[63] An assassination is defined as any politically motivated murder or attempted murder of a high government official or politician. A revolution is any illegal or forced change in the top governmental elite, any attempt at such a change, or any successful or unsuccessful armed rebellion whose aim is independence from the central government.

imputed so that the international diffusion measures cover the whole study period 1952-1985.

First we must clarify the definitions of the BOF and open hearth as the alternative methods of production. What we refer to as the open hearth output is in fact a joint category of old technology use, that is, it includes output produced using both open hearth and Bessemer process. The Bessemer process is the oldest of the technologies and its proportion is small in our sample, except in Belgium. If Bessemer output was excluded, the open hearth data could not on its own be interpreted as a measure the displacement of old technology by the BOF. The electric arc in contrast should not be included because at the time, it was used mainly for the production of small-scale specialised steel and so cannot be considered a substitute (Rosegger 1980). By the basic oxygen process we also mean output produced using the Kaldo and Rotor processes, which also use oxygen.[64] This is consistent with the way output is broken down by process in the OECD publication "The Iron and Steel Industry", which is the source of data for HCCTAD. Meyer and Herregat (1974: 152 footnote) state that Kaldo and Rotor were not successful in terms of diffusion but are included in the figures "to be complete".

The coverage of the data in HCCTAD is generally good and the data is of good quality. Missing values were added and some corrections made after consulting the original source, various issues of "The Iron and Steel Industry" by the OECD, and also some issues of "The European Steel market" by the UN Economic Commission for Europe. We imputed zero values for BOF capacity from 1952 if it was clear that

[64] Kaldo/Rotor is used in Sweden.

the first observed BOF tonnage correctly represents the first adoption date in that country. If the first reported tonnage was very small, this was taken as confirmation that the technology was not previously used, and otherwise we again referred to commentary in the OECD publication to confirm that production had indeed started in that particular year. These zero value imputations are important because the sample size increases considerably for the dependent variables BOF capacity and BOF proportion. For example, for Australia we impute zeros for 1952-61 which increases the sample size by 10 observations although tonnage breakdown is missing until 1968.

For the purposes of the international diffusion measures, the following imputations were also made for four countries. For the United States and Japan, only total output is available in 1952-3. We used the breakdown by process in 1954 to approximate the proportions in 1952 and 1953, allowing for some change according to the observed time trends. For Canada and Australia, we used other sources to obtain breakdown by process for the years before full data is available; 1963 (Canada) and 1968 (Australia). We have the proportion of electric arc output in 1961-2. For the earlier period we also have breakdown by process in 1954 for Canada, which is also the year in which basic oxygen furnaces were first installed (Singer 1969:4).[65] We then used the method of random imputations, and ordered the numbers according to the apparent time trend in the shares of each process. This procedure consists of assigning a random number between the two closest observed values to the missing year, and if two observations are missing, the random numbers are ordered if there is an obvious time trend during the period. This is the same

[65] See also www.dofasco.ca.

procedure as in Chapter 3 and we also use it for years of schooling and political instability in this chapter (section 5.3.3.3). Finally, in the case of Australia BOF production began in 1962 but the OECD does not provide a breakdown by process until 1968. From O'Malley (2000:723) and Richards (2004) we obtain a figure for BOF capacity installed in 1962. We then impute the missing values allowing again for randomness in both open hearth and BOF output. These imputations were only used for A and A%, not the dependent variables (intra-country diffusion).

The theoretical model refers to capacity which is much less sensitive to demand shocks than output is. However, capacity data is rarely available. In the "The Iron and Steel Industry", the level of precision with which capacity data is reported is too low to use for estimation. In her study Oster (1982) consults annual reports as well as directories of the American Iron and Steel Institute in order to obtain figures of capacity for United States firms. Faced with this difficulty, we choose to use the very good data on output. We smooth the data using the assumption that annual fluctuations in demand affect output but not capacity. In particular, we smooth away downward movements in BOF output before 1974 and upwards movements in open hearth output. The resulting measure we refer to as "capacity". It is below true capacity because steel is rarely produced at full capacity. However, if the distortion – the amount by which our measure underestimates capacity – is constant across countries, which we consider a reasonable assumption, then the fact that true capacity is higher does not bias our estimates.

The precise process we use to smooth the data is the following. First, we identified the peaks in open hearth and basic oxygen output in the sample period. Before the peak, capacity is computed as the maximum output between 1952 and the current period. For basic oxygen we have:

$$CBOF_{Ct} = \max\left[BOF_{C,t=0},\dots,BOF_{Ct})\right] \tag{5.13}$$

for all years t until the observed output peak, and similarly for open hearth capacity. The open hearth output peak in our sample occurs between 1953 (Austria) and 1968 (Canada). In each following year, open hearth capacity is computed as the maximum output observed between the current period and the last observed period T when open hearth capacity is zero:

$$COH_{jt} = \max\left[OH_{Ct},\dots,OH_{CT})\right]. \tag{5.14}$$

In the case of basic oxygen our assumption is that in the absence of change in demand, capacity continues to increase until the displacement of open hearth furnaces is completed. This is what we would expect to see, on the basis of the theoretical model alone. However, as discussed above, there was a considerable fall in demand in the mid-1970s and subsequently capacity was cut. These cuts are not part of the diffusion process we are interested in, but we must take them into account. We took a cautious approach and only allowed a decrease in capacity if there is a considerable and consistent fall in output so that it does not recover to the peak level. For the years following the BOF peak, capacity is given by

$$CBOF_{Ct} = \max\left[BOF_{C,PEAK},\dots,BOF_{CT})\right] \tag{5.15}$$

until the end of the study period. The countries for which we allow capacity to fall in this way are: Belgium (peak in 1974), France (1980) Germany (1974), Italy (1981), Japan (1973), Luxembourg (1979), Spain (1976), United States (1979).

The capacity measure, constructed as explained here, is used in computing each of the four dependent variables and the two measures of international diffusion. The proportion of basic oxygen in each country is computed as $CBOF_{Ct}/(CBOF_{Ct}+COH_{Ct})$. The total basic oxygen capacity in the world is computed as the unweigthed sum over all 15 countries in the sample: $A_t = \sum_{C=1}^{15} CBOF_{Ct}$. The world BOF proportion is the simple average, total BOF capacity divided by total capacity: $A\%_t = \sum_{C=1}^{15} CBOF_{Ct} \div \sum_{C=1}^{15} (CBOF_{Ct} + COH_{Ct})$. The regressors A and A% used in the empirical model are measures of the extent of use elsewhere. These are country-specific and obtained by subtracting the contribution of the relevant country from the measure of overall international diffusion.

Table 5.1 reports the number of observations, minimum, mean, and maximum values and the standard deviation for each diffusion measure and measures A and A% computed over all 15 countries. Annual change in the international diffusion measures is also reported which is used in the discussion of the empirical results.

Table 5.1 Means, maximum and minimum values

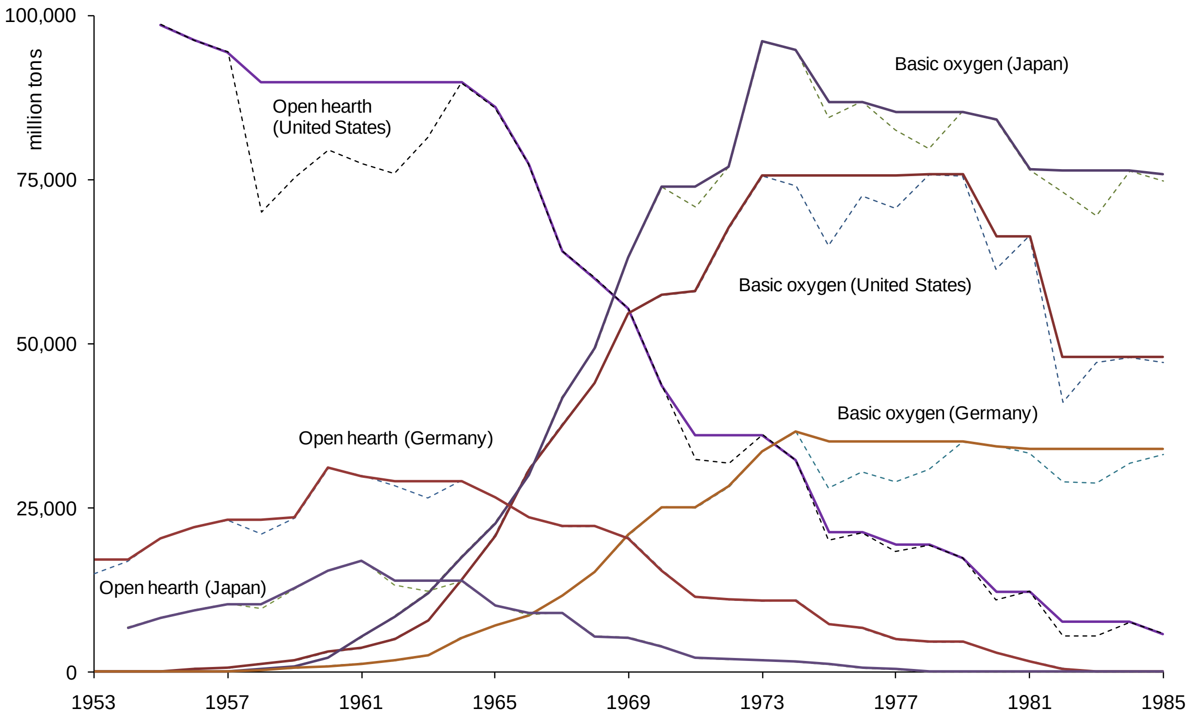
Variable	# obs	# missing	min.	mean	max.	std. dev.
ΔCBOF / COH	379	131	0	0.29	28.9	1.62
-ΔCOH / COH	291	219	0	0.18	1	0.27
BOF share	493	17	0	0.51	1	0.39
BOF capacity ('000 tons)	493	17	0	10,060	96,100	19,000
A%	510	0	0	0.48	1	0.38
A ('000 tons)	510	0	0	132,000	289,000	108,000
ΔA%	495	15	0	0.030	0.087	0.025
ΔA ('000 tons)	495	15	-43,700	6,870	43,700	17,100

Notes to Table: If two outliers are ignored, the mean of ΔCBOF/COH is 0.18, the maximum is 3.56, and the standard deviation is 0.45.

Availability of data on the explanatory variables limits the sample to 13 countries in the period 1960-1985. When either ΔCBOF/COH or -ΔCOH/COH is used as the dependent variable, the sample is smaller because data is only used until the year that open hearth output ceases in each country. In the case of -ΔCOH/COH we take as the first observation the open hearth output peak or the first observed BOF tonnage, whichever is later, because conceptually this measure requires that open hearth output is falling and BOF output is rising.

The unbalanced nature of the panel is not problematic as long as the structure of unbalancedness is exogenous, that is, there is no selection bias if selection does not depend on the endogenous variables (Verbeek 2000:217, Hsiao 1986). Canada and Australia enter the sample later, however for reasons quite separate from the steel sector these countries are not included in the list of countries in the OECD publication which is the source of our data. Unbalancedness creates computational issues but the software we use (PcGive) is designed to handle these. We also have no problem with the minimum number of consecutive time-periods that is available, since the smallest number is already large.

Figure 5.1. Basic oxygen and open hearth output and capacity (Germany, Japan, United States) 1953-1985

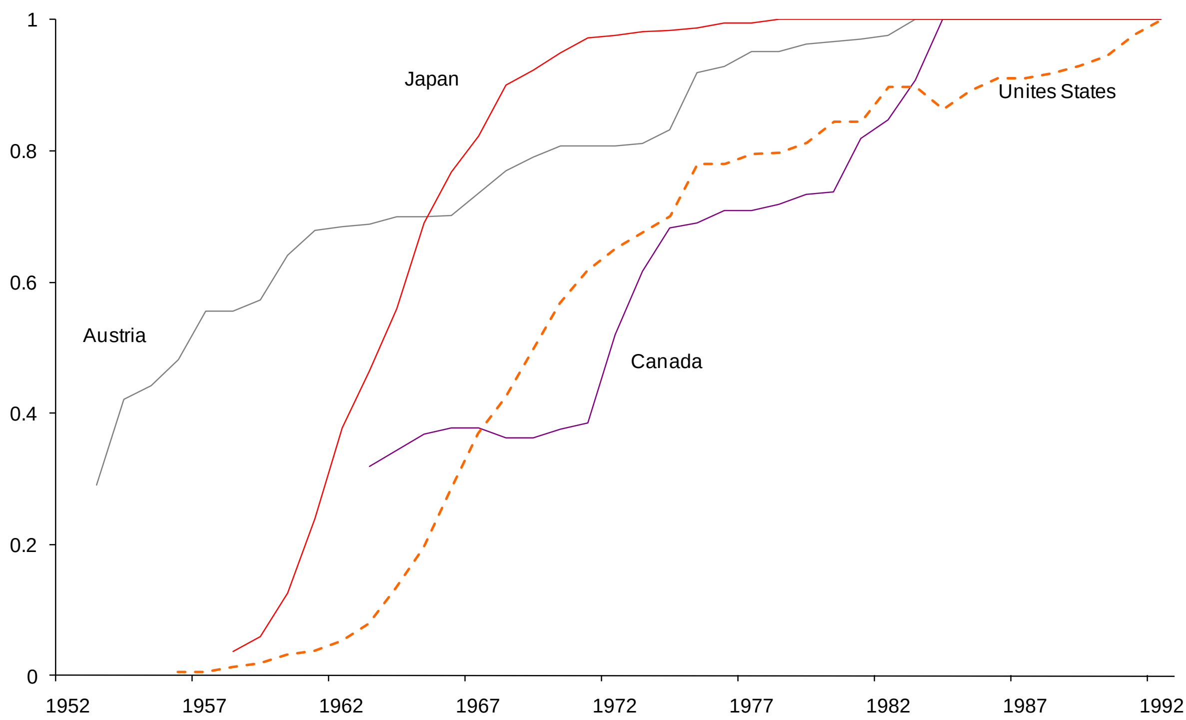


Notes: The dotted lines are the original data (output) and the solid lines are the capacity measures we have constructed.

We plot BOF and open hearth output and capacity for a selection of countries in Figure 5.1. The dotted lines indicate the output measure while the solid lines indicate our capacity measure. We see that although the first use of BOF took place in 1953 (in Austria), in other countries output only begins to increase strongly in the 1960s. Output increases until the mid-1970s after which it there are some drops in output (Japan, United States) or otherwise tonnage is relatively constant until the end of the study period (Germany). Open hearth output grows for a number of years after 1952. This is contradictory to the assumption that because the open hearth is inferior to BOF, no new investment in the open hearth takes place once the BOF becomes available. In our theoretical model the period between 1952 and the open hearth peak is attributed to an increase in the demand for steel. For the dependent variable -ΔCOH/COH we only use the data following the open hearth peak, which, as indicated here, occurs for most countries before the “take-off” of BOF in that country.

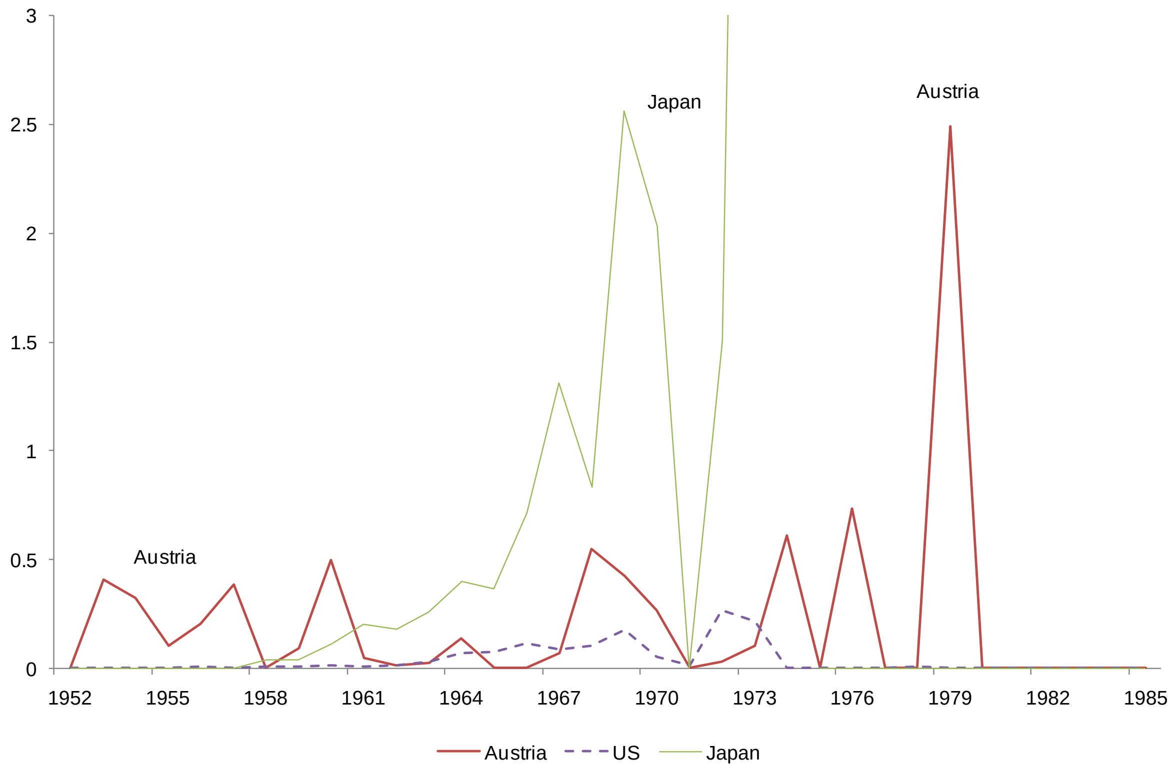
The BOF share is plotted for a selection of countries in Figure 5.2. The time paths have some of the S-shape that is typically expected of a diffusion curve. However, the patterns are by no means smooth and there are differences in the start date of diffusion and the steepness of the path. Whereas the proportion reaches unity for all other countries by 1985, this only occurs in the United States in 1992. Austria, where the BOF was invented, is distinctive among all the countries as diffusion starts early and at a high level. Austria is also an example of a country in which the BOF proportion increases relatively slowly.

Figure 5.2. Basic oxygen proportion, 1952-1992 (Austria, Canada, Japan, United States)



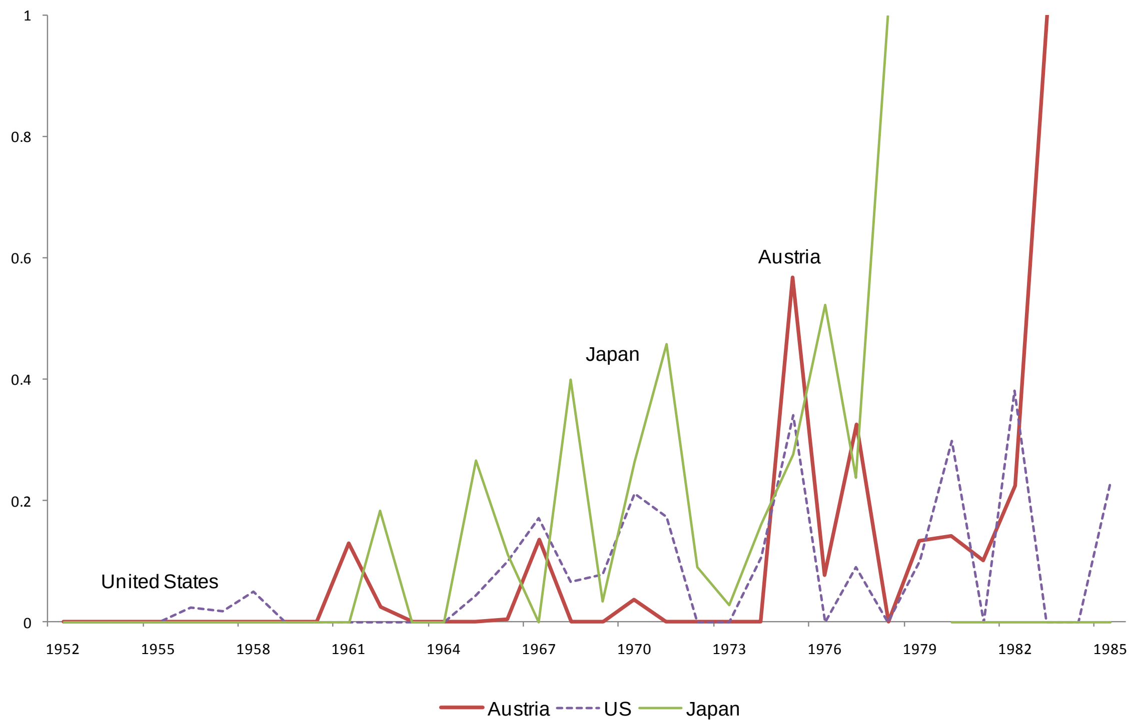
In Figure 5.3 we plot the measure ΔCBOF/COH. There is no clear time trend in this measure, and the magnitude can be very large or, as for the United States, relatively small throughout the period. Typically, at least one large value is observed which corresponds to the year in which the open hearth is last used. The last observed value (1973) for Japan is off the scale (10) for this reason. This is the second highest value in our sample and we include a dummy variable to control for it.

Figure 5.3. Rate at which BOF capacity increases relative to remaining open hearth capacity, ΔCBOF/COH, 1952-1985 (Austria, Japan, United States)



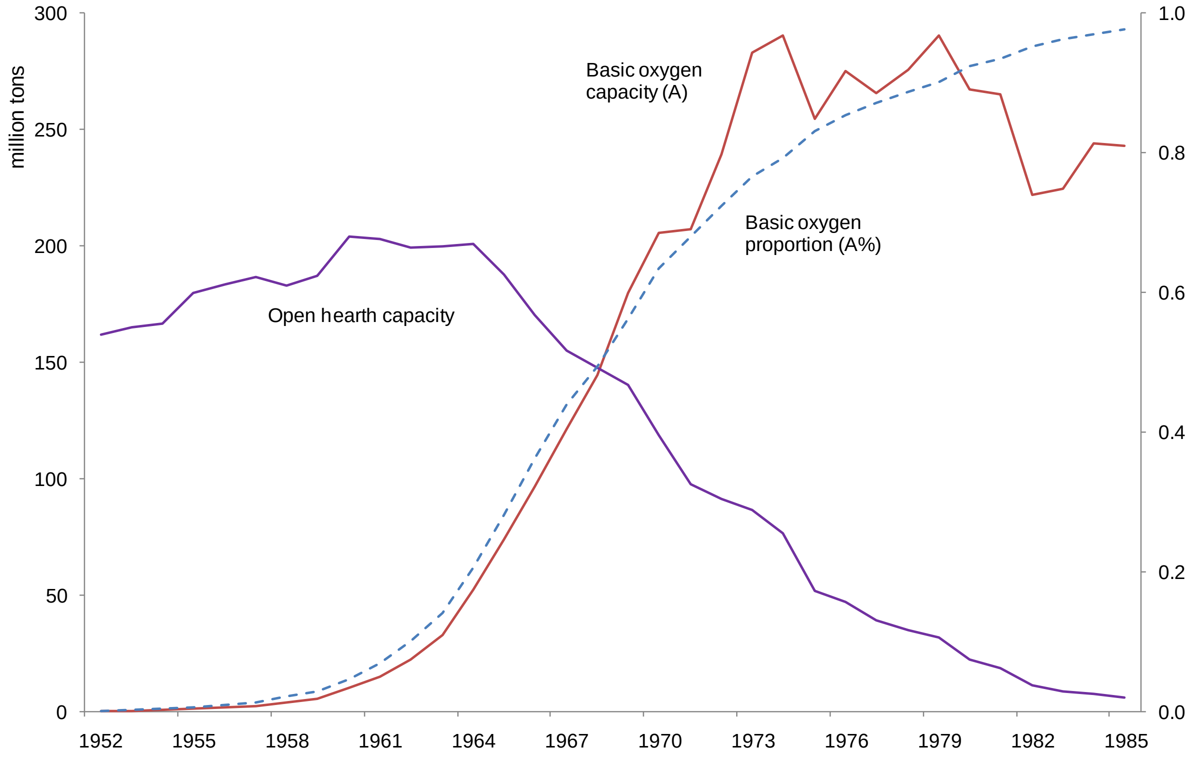
The variable -ΔCOH/COH is plotted for three countries in Figure 5.4. It lies between zero and unity in all periods and equals unity in the last year in which open hearths are used. The differences in estimation period across the two measures of change can be illustrated by comparing the last observed values for Japan. Open hearth production continues in Japan until 1977 so the measure -ΔCOH/COH equals unity in 1978. However, the measure ΔCBOF/COH is not defined after 1973 because BOF capacity begins to fall. This implies that in the case of Japan, the sample period ends earlier when ΔCBOF/COH is used as the dependent variable compared to the other dependent variables.

Figure 5.4. Rate at which open hearth capacity falls, -ΔCOH/COH, 1952-1985 (Austria, Japan, United States)



Finally, the international diffusion measures together with aggregate open hearth capacity are plotted in Figure 5.5. Our measure of international BOF capacity peaks at 290,500 million tons in 1974, when the share of BOF is 79 per cent. In subsequent years the proportion continues to increase but capacity does not, which reflects the fact that BOF capacity is constant in many countries and falling in several (see above). The BOF proportion is a very smooth S-shaped curve. It reaches unity in 1992 as the last open hearths disappear from the United States.

Figure 5.5. International BOF capacity and proportion, open hearth capacity, 1952-1985



5.5 Results

5.5.1 Overview

Estimates of the model are presented in Table 5.2. In the first three columns the dependent variable is a measure of change: change in BOF capacity relative to remaining open hearth capacity (columns 1 and 2), and negative growth rate of open hearth capacity (column 3). In columns 4-6 we use level measures of the extent of use: BOF share in column 4, and BOF capacity in columns 5 and 6. A dummy is included in columns 1 and 2 for the highest observed value of the dependent variable: ΔCBOF/COH is 10 in Japan in 1973[66] whereas the next highest observed value is

[66] The value is 29 in Luxembourg in 1978 but this country is not included because data on the schooling measures is not available.

only 3.6. The value for Japan is so high because the denominator is the last observed open hearth tonnage which is very small.

The extent of diffusion elsewhere is measured either as BOF capacity in other countries (A) or the share of basic oxygen in total capacity in other countries (A%). Our theoretical model does not tell which measure should be used in each case so we use the following criteria. For the level measures of diffusion the choice is determined by which variable makes the results more straightforward to interpret so A ('000 tons) is used for BOF level ('000 tons), and A% is used for BOF proportion (both measured as proportion of total tonnage). For the change measures of diffusion we estimate the models using both A and A% and then report the results of the model which fits the data better. Thus for -ΔCOH/COH we use A (million tons) and for ΔCBOF/COH we use A%.[67]

The other regressor that varies somewhat across the different specifications is the time control term. We present the results using the nonlinear term $K=2^{-0.1t}$ but also biannual time dummies in columns 2 and 6. The motivation for column 2 is that the coefficient on K in column 1 has the wrong sign. Using BOF capacity as the dependent variable results are not at all robust to the choice of time control and this is why we report results with time dummies in column 6. The results in column 5 are suspect because they are quite different from results in other columns and not robust for example to the inclusion of ΔK.

[67] To evaluate model fit we use equation standard error. With ΔBOF/COH equation standard error is 0.504 using A% and 0.514 using A (column 1), and 0.472 using A% and 0.485 using A (column 2). With -ΔCOH/COH as the dependent variable, equation standard error is 0.239 using A and 0.242 using A%.

Table 5.2 General model

Explanatory variables	Dependent variable (S): ΔCBOF / COH		- ΔCOH / COH	BOF share	BOF capacity	
	1	2	3	4	5	6
S (t-1)	-0.105 * (0.0604)	-0.214 ** (0.0673)	0.261 ** (0.116)	0.915 ** (0.0423)	0.853 ** (0.0597)	0.81 ** (0.0456)
A% (t-1)	-1.93 ** (0.378)	-12.1 ** (5.55)	-	-0.113 * (0.0607)	-	-
A (t-1)	-	-	-0.00119 ** (5.88E-04)	-	0.00965 ** (0.00485)	-0.039 ** (0.0175)
ln world GDP	9.18 ** (2.33)	7.03 ** (2.14)	-2.8 ** (0.953)	0.0267 (0.167)	176 ** (43.6)	30,800 * (16,900)
primary years	0.024 (0.25)	0.147 (0.163)	0.182 (0.138)	0.0479 ** (0.0204)	-465 ** (114)	3,350 (3,420)
secondary years	-0.19 (0.161)	0.192 * (0.109)	-0.257 ** (0.0527)	-0.00747 (0.0134)	-5,120 ** (1,260)	-5,550 ** (1,460)
pol. instability	-0.882 (1.29)	-0.718 (1.38)	-0.0451 (0.63)	0.183 ** (0.0787)	-96.0 ** (23.6)	7,270 * (3,940)
nominal r	3.16 (3.42)	6.44 (5.27)	-3.05 (2.58)	0.281 ** (0.143)	-47.7 ** (11.7)	-7,330 (9,110)
CPI Δ%	-1.35 (2.23)	-1.26 (2.15)	0.585 (0.87)	0.163 * (0.102)	-41.6 ** (10.2)	32,800 ** (7,570)
exchange rate	-1.47E-03 * (7.78E-04)	-6.39E-04 (8.14E-04)	2.18E-04 (5.26E-04)	-2.54E-06 (2.8E-05)	5.1 (5.52)	4.1 (4.26)
Δe	0.00118 (9.05E-04)	0.00115 (8.45E-04)	8.80E-04 (5.68E-04)	5.37E-07 (4.75E-05)	-12.2 (12.4)	-3.28 (6.23)
$K=2^{-0.1t}$	11.4 ** (2.54)	no	-1.87 (1.83)	-0.732 ** (0.271)	-159 ** (39.3)	no
ΔK	-0.817 ** (0.182)	no	0.135 (0.132)	0.0525 ** (0.0194)	11.4 ** (2.82)	no
time dummies	no	yes	no	no	no	yes
constant	-0.0378 (0.0695)	1.19 ** (0.279)	0.133 ** (0.0202)	-0.00743 ** (0.00232)	193 ** (74.8)	1.22 (557)
No. of obs.	201	201	212	271	271	271
No. of countries	12	12	13	13	13	13
AR(1)	-2.05 **	-2.39 **	-2.98 **	-2.66 **	-1.34	-1.37
AR(2)	-0.432	-0.974	0.828	-1.48	-0.715	-0.473
Sargan test	214 * (190)	166 (190)	205 (190)	235 ** (190)	301 ** (190)	248 ** (190)
eq. st.error	0.504	0.472	0.239	0.0381	2904	2850
Wald (S,A)	31 ** (2)	12 ** (2)	6.3 ** (2)	625 ** (2)	371 ** (2)	333 ** (2)
Wald (8)	110 ** (8)	31 ** (8)	100 ** (8)	35 ** (8)	17 ** (8)	39 ** (8)
Wald (edu,pol)	3.5 (3)	11 ** (3)	48 ** (3)	6.8 * (3)	17 ** (3)	23 ** (3)
Wald (r,CPI)	1.0 (2)	1.6 (2)	1.4 (2)	12 ** (2)	17 ** (2)	19 ** (2)

Notes: Standard errors in parentheses. * indicates significant at 10 per cent, ** at 5 percent. A is measured in 1M tons in column 3 and in 1000 tons in columns 5 and 6. BOF capacity is in 1000 tons (columns 5 and 6). Dependent variables in 1-4 are expressed as shares.

In short, there is evidence that the lagged extent of use (domestic and foreign together) is a significant determinant of current domestic diffusion. The international stock effect is negative but the domestic stock effect is positive. The eight other explanatory variables – world GDP, years of primary and secondary schooling, political instability, nominal interest rate and CPI change, exchange rate and annual change in exchange rate – are also jointly significant in all columns. The test statistic is reported on the row "Wald (8)". Few of the variables are individually significant. Estimates of world GDP are most robust as this is significant with the expected positive sign with two of the four dependent variables (columns 1, 2, 5 and 6). Years of schooling and political instability are jointly significant at least at 10 per cent in all columns except column 1 ("Wald(edu, pol)"). The nominal interest rate and change in CPI, which together measure the real interest rate effect, are also jointly significant in columns 4-6 ("Wald (r,CPI)"). Where the time control K has a negative coefficient, the change in K has a positive coefficient as expected. We discuss the results in detail, beginning with the national and international stock effects S(t-1) and A(t-1) or A%(t-1).

First, however, we consider the results of the specification tests proposed by Arellano and Bond (1991) are reported. The AR(1) and AR(2) statistics test for the absence of serial correlation in the disturbances of the levels equation. Consistency of GMM requires that e_t and e_{t-1} are not correlated; in the differenced equation, this means that there should be no evidence of second-order serial correlation i.e. correlation between $\Delta\varepsilon_t$ and $\Delta\varepsilon_{t-2}$ (ε_t-ε_{t-1} and ε_{t-2}-ε_{t-3}). Significant negative first-order correlation is expected. The other specification test is the Sargan test of overidentifying restrictions. The null hypothesis is that all the GMM instruments are valid; the degrees of freedom is equal to the number of overidentifying restrictions under the null hypothesis.

There is no evidence of second-order serial correlation in any of the regressions. This indicates that errors in the levels equation are not correlated and that GMM is consistent. The AR(1) statistic is negative as expected however not significant in columns 5 and 6 which is surprising given that $\Delta\varepsilon_t$ and $\Delta\varepsilon_{t-1}$ share the term ε_{t-1}. The Sargan test for the validity of GMM instruments is passed in columns 1-3 (at 10 per cent in column 1) but not in columns 4-6. This is likely to be at least partly due to heteroskedasticity because the Sargan statistic for one-step GMM estimates is not heteroskedasticity-consistent: the test statistic only has a limiting chi-squared distribution if the errors are IID across countries and time (Arellano and Bond 1991). Although the diagnostic test results are not fully satisfactory, we are reasonably confident about the consistence of GMM because the AR(2) tests are passed in all columns.

A serial correlation test is also a test of whether there is evidence against parameter homogeneity in the sample, that is, evidence that the slope parameters are not the same across countries (Durlauf et al. 2005). For example, if the coefficient on $A_{-i,t-1}$ was different across countries then the error term would contain an element resembling $\left(\hat{\beta}_{2i} - \beta_2\right)A_{-i,t-1}$ which creates serial correlation in the errors because $A_{-i,t-1}$ is persistent. The fact that the AR(2) test is passed in each regression suggests therefore that the assumption of parameter homogeneity, which we have implicitly made, is not violated.

5.5.2 Stock effects

The lagged dependent variable and the extent of use elsewhere are jointly significant in all columns (see row “Wald S,A”). Both variables are also individually significant at least at 10 per cent. Two questions arise: Does past domestic diffusion have the same effect as past diffusion elsewhere? And is the magnitude of the international effect economically significant? When domestic and foreign past diffusion are measured in the same way – A and BOF capacity (thousand tons) in columns 5 and 6, and A% and BOF proportion (of all capacity) in column 4 – the first question can be answered by testing whether the coefficients on S_{t-1} and $A_{-i,t-1}$ are significantly different from each other. We use a χ^2 test of the linear restriction that the effects are equal and find that the null is rejected in columns 4-6 (test statistics 327, 181 and 245). This means that the location of past diffusion matters, that is, the domestic and international stock effects are not equal in magnitude.

The results also suggest that the two effects work in opposite directions which contrasts the hypothesis that both stock effects are negative. The expected negative coefficient on the lagged dependent variable is only obtained in columns 1 and 2. There is strong support for the hypothesis of a negative international stock effect however: the extent of use elsewhere has a negative coefficient except in column 5 and this is significant at 5 per cent in all columns except in column 4 where the p-value is 0.064. These results suggest that there is a negative international stock effect but that past domestic diffusion has a positive effect on current diffusion.

The regression in column 5 only differs from column 6 in the choice of time control and it is clear that most coefficient estimates are not robust to that choice. Use elsewhere is statistically significant in both columns but with the unexpected positive sign in column 5. The estimate of past domestic diffusion is robust with a similar point estimate and standard error in both columns; it is positive and significant as in columns 3 and 4.

The coefficient on the lagged dependent variable is between 0.81 and 0.915 in columns 4-6 which use a level measure of diffusion. The coefficient is considerably smaller in columns 1-3 which possibly reflects difference in persistence; the stock measures are more persistent over time (see plots in section 5.4). Since our measure of capacity is derived from output it is likely that actual capacity has an even higher degree of persistence than our measure. Interestingly the estimates in columns 1 and 2 suggest that the international effect is stronger than the domestic effect. The negative coefficient on the lagged dependent variable in these columns is what we expected based on the hypothesis that the stock effect is determined by the extent of international diffusion and thus past use whether at home or abroad has a negative effect on further use. We discuss possible reasons for the positive coefficient in other columns in section 5.7. The estimates of the lagged dependent variable also determine the difference between the short- and long-run estimated effects of the other variables. The positive coefficients in columns 3-6 imply that the long-run coefficients in these regressions exceed the short-run effects whereas in columns 1 and 2 the short-run effects are larger.

To evaluate the economic significance of the international effect, we examine the expected effect of an average annual change in use elsewhere on domestic diffusion. In Table 5.1, mean values of each of the dependent variables are reported together with the mean annual change in A and A%. On average, the extent of basic oxygen capacity elsewhere increases by 6.9M tons (A) or 3.0 percentage points (A%)

annually. The mean hides considerable differences within the period: the median increase is only 4.8M tons or 2.1 percentage points. The average annual increase was as high as 12M tons or 3.6 percentage points over 1952-1974 with a median of 6.9M tons (i.e. equal to the average over the whole period) and 3.7 percentage points. After 1974 although BOF share continued to increase capacity did not (see Figure 5.5 in section 5.4). Because tonnage is more volatile than the share, the averages of A and A% do not correspond to each other very well, however, for the purposes of this exercise, we have considered the mean changes in A% and A to be comparable.

We compute both short- and long-run estimated effects together with 95 per cent confidence intervals. Long-run estimates are presented in Table 5.3; these can be obtained by diving the short-run coefficient by 1 minus the coefficient on the lagged dependent variable. Consider first BOF share and BOF capacity (columns 4-6). A 3 percentage point increase in A% is estimated to reduce the domestic share by 0.34 percentage points (confidence interval (-0.70, +0.02) points) in the short run and by 4.0 (-8.3, +0.31) points in the long run, ceteris paribus. A 7M ton increase in A is estimated to reduce the domestic BOF tonnage by 270,000 tons (-510,000, -33,000) in the short run and by 1.4M tons (-3.1M, +0.18M) in the long run (column 6), ceteris paribus. The estimated effect in column 5 is an increase of 68,000 tons (+1, +130,000) in the short run and of 460,000 tons (+81,000 +840,000) in the long-run.

Table 5.3 Long-run coefficient estimates and standard errors

	Dependent variable					
Explanatory	ΔBOF / COH		- ΔCOH / COH	BOF share	BOF capacity	
variables	1	2	3	4	5	6
A% (t-1)	-1.75 (0.362)	-9.95 (4.42)		-1.33 (0.732)		
A (t-1)			-0.00161 (9.4E-04)		0.0658 (0.0277)	-0.205 (0.118)
ln world GDP	8.31 (2.22)	5.79 (1.77)	-3.79 (1.56)		1,200 (714)	162,000 (68,900)
primary years				0.563 (0.415)	-3,170 (1,880)	
secondary years		0.158 (0.0878)	-0.334 (0.0832)		-34,900 (20,700)	-29,200 (12,700)
pol. instability				2.16 (1.79)	-655 (387)	38,200 (19,900)
nominal r				3.31 (2.02)	-325 (192)	
CPI % change				1.92 (1.09)	-283 (168)	172,000 (57,600)
exhange rate	-0.00133 (7.31E-04)					

Notes: Long-run standard errors in parentheses.

Because the effects are not precisely estimated, that is, the confidence intervals are rather wide, we cannot make strong statements about the economic significance of the international effect. The boundaries of the intervals indicate that the magnitude can range from very large to so close to zero that it is economically uninteresting; and a positive long-run effect cannot be ruled out even in regressions where the point estimate is significantly negative. Interestingly, the point estimate from column 6 suggests that the international effect is so large that, ceteris paribus, an increase in the extent of use in one country would lead to an overall fall in world diffusion. To illustrate this point, consider a world which consists of seven countries. If all the 7M ton increase occurs in one country, capacity in each of the other six countries falls by 1.4M tons in the long run which, ceteris paribus, implies an overall reduction in international diffusion of 1.4M tons. The estimated short-run overall effect is that international increases by 5.4M tons (7M-6*0.27M) which also indicates a considerable magnitude for the international effect. Of course in practice other factors are simultaneously encouraging further diffusion and so we would not for example expect to see an actual fall in diffusion.

Diffusion is measured as a tonnage in column 6 and as a proportion in column 4, but using the sample averages we compare the estimated magnitudes of the international effect. The procedure we use requires the assumption that a lower domestic BOF output implies that domestic open hearth output is higher by the same amount so that total output is constant. This is not exactly what the theoretical model implies but it facilitates the comparison of percentages and tonnage levels in this exercise.[68] Suppose that domestic diffusion is initially at the sample average, which is 50.6 per cent or 10.06M tons. The implied total tonnage is 19.88M. Then:

The estimated ceteris paribus effect of a 7M ton increase in A is that domestic tonnage falls to 9.79M (column 6, reduction of 273,000 tons)[69] in the short run. Assuming that total tonnage is constant at 19.88M (i.e. open hearth output is 273,000 tons higher), the new implied BOF share is 49.2 per cent.

A 3 percentage point increase in A% is estimated to reduce the domestic share by 0.339 percentage points in the short run which implies that the new share is 50.26 per cent (column 4). The new share implies that the new domestic BOF tonnage is 9.99M tons, assuming that total tonnage is constant at 19.88M.

Thus the BOF share regression suggests a short-run effect which is only one quarter of the BOF capacity estimate (70,000 tons vs. 273,000 tons; or 0.34 percentage points vs. 1.4 percentage points). The long-run estimates are closer to each other; the BOF capacity estimate is less than twice as large as the BOF share estimate. Assuming again the same initial values, we have:

---

[68] Because each open hearth producer produces less than each BOF producer, total output should be smaller by the amount that reflects this difference. See section 4.3.
[69] The estimate is the same as was referred to earlier; here we use one more significant digit.

A 7M ton increase in A reduces domestic tonnage to 8.63M (column 6, reduction of 1.44M tons) in the long run. Assuming that total tonnage is constant at 19.88M, the new implied BOF share is 43.4 per cent (column 6; reduction of 6.66 percentage points).

A 3 percentage point increase in A% is estimated to reduce the domestic share by 3.99 percentage points in the long run which implies that the new share is 46.6 per cent (column 4). The new share implies that the new domestic BOF tonnage is 9.27M tons (reduction of 0.79M tons), assuming that total tonnage is constant at 19.88M.

This comparison suggests to us that at least at the sample mean level of diffusion, the estimates of not only the sign but also (with some reservations) the magnitude of the international effect are robust to the choice between two measures of the extent of use, basic oxygen capacity and proportion. This conclusion is based on comparing point estimates. In the theoretical model total capacity is not constant as we assumed here, but expected to be somewhat lower if BOF tonnage is lower (depending on the difference between domestic BOF and open hearth producers' outputs). Taking this into account brings the estimates from columns 4 and 6 even closer to each other.[70] It can also be noted that in the exercise we just described, if we compare short- and long-run confidence intervals rather than only point

[70] Suppose that when BOF output falls, total output falls by 25 per cent of this amount. Then, the estimate from the BOF capacity regression is that the new BOF proportion is 49.4 in the short run and 44.2 per cent in the long run. These are higher than if we assume that total tonnage is constant. In contrast the BOF share estimates of the tonnage reduction are higher if we assume that total tonnage falls. This brings the two sets of estimates closer together however any assumption about how much total tonnage falls is arbitrary, not based on data or other information.

estimates then we find that these overlap.[71] The areas of overlap suggest that an average annual increase in international diffusion reduces the domestic extent of use (from the sample average level) by up to 510,000 tons in the short run and by 1.65M tons in the long-run.

The fact that the estimated international effect in column 6 is comparable to that from column 4 suggests to us that the column 6 estimates are in this sense preferable to those from column 5. When we later examine other coefficient estimates this conclusion is supported.

We now turn to the estimates of the international effect in columns 1-3. We refer to ΔCBOF/COH, the change in BOF capacity as a share of remaining open hearth capacity, as the relative rate of growth and -ΔCOH/COH we call the rate of switching to indicate that producers replace open hearths by the basic oxygen furnace. As before, 95 per cent confidence intervals are presented in brackets. A 3 percentage point increase in A% reduces the relative rate of growth by 5.8 (-8.0, -3.6) percentage points in the short run and by 5.2 (-7.4, -3.1) percentage points in the long run (column 1), ceteris paribus. Relative to the sample average value of ΔCBOF/COH, 18 per cent,[72] the estimate of the international effect is of a considerable magnitude both in the short and long run. The long-run estimate is smaller than the short-run because the coefficient on the lagged dependent variable is negative. Using this dependent variable, the use of time dummies instead of K in column 2 has the unfortunate consequence that the estimate is very imprecise and of

---

[71] The 95 confidence intervals for the new BOF tonnage are (9.92, 10.06) and (9.55, 10.03) for the short-run effects and (8.41, 10.12), and (7.0, 10.24) for the long-run effects (columns 4 and 6 respectively). This assumes that total tonnage is constant.

[72] The average in the full sample is 0.29 but excluding the two highest values gives the average value of 0.18. See Table 5.1.

a magnitude which is not economically sensible. The estimated ceteris paribus reduction is 36 (-69, -3.7) percentage points in the short run and 30 (-56, -3.9) percentage points in the long run (column 2). The upper bound of the confidence interval is similar to the lower bound of column 1 estimates and this is the magnitude that is economically most sensible, although still perhaps surprisingly large. The impact of the time dummies is discussed in some more detail below.

In contrast, results in column 3 suggest a rather small international effect. A 7M ton increase in A is estimated to reduce the rate of switching by 0.83 (-1.6, -0.026) percentage points in the short run and by 1.1 (-2.4, +0.16) percentage points in the long run, ceteris paribus. Given that the dependent variable is between 0 and 100 per cent, these magnitudes are small. However, the results add to the evidence for a negative international effect which we have found across all four measures of diffusion. Using either dependent variable which measures change in diffusion the difference between the short and long run effects is quite small; most of the international effect happens immediately. In contrast our results suggest that the level measure are persistent.

In the conclusion to this Chapter (section 5.7) we discuss the finding that the lagged dependent variable has a positive coefficient in more detail. This result contradicts the theoretical hypothesis of a stock effect in the sense that, in theory, both domestic and foreign past extent of use should have a negative effect on current domestic use. The reason is that the profitability of adoption depends on the total extent of use in the world. First however we turn to the results regarding the eight other explanatory variables: human capital and political instability, world real GDP, the real interest rate, and the exchange rate.

5.5.3 Human capital and political instability

As a group, the two measures of schooling and political instability are statistically significant at 10 per cent or less in all columns except column 1 ("Wald (edu,pol)"). Years of primary and secondary school education are also jointly significant in columns 3 to 6. Each variable is also individually significant at 5 per cent at least in two columns but many of the signs are opposite to what we expected. Recall that we expected schooling to have a positive effect and political instability to have a negative effect on diffusion. A one year increase in primary schooling is estimated to increase BOF share by 4.8 percentage points (56 points) in the short (long) run but to reduce BOF capacity by 465,000 tons (3.2M tons) in the short (long) run. A year of secondary schooling is estimated to increase ΔCBOF/COH by 19 percentage points (16 points) in the short (long) run but to reduce -ΔCOH/COH by 26 points (33 points) and BOF capacity by over 5M tons (29M tons) in the short (long) run. These point estimates suggest that secondary schooling has a greater effect on diffusion but the magnitude of both effects is very large.

Political instability is estimated to have a negative effect on BOF share and also a significant effect on BOF capacity however the sign is not robust to the time controls. The positive sign is difficult to explain, but the expected negative coefficient was rarely obtained across a number of different specifications which we estimated. An explanation may lie in the lack of variation in this measure: it is mostly zero or very small. The variation that exists is dominated by a few relatively large (positive) values namely in the United Kingdom in 1970-84, Italy in 1975-9, Germany in 1980-4, and Spain in 1975-84. It may be that the variable acts as a kind of country fixed effect and that it reflects some other factor common to these countries and periods which had a positive effect on diffusion. It may also be that time-series variation is unimportant because the variable that matters is the long-term underlying stability of the political system; if so, this can only be captured if there is sufficient cross-sectional variation that it we would need a larger sample of countries. In this data there is no evidence that political instability has the hypothesised negative effect on diffusion.

The evidence of a human capital effect is worth some more investigation. A possible explanation for the lack of robustness is that years of schooling (divided into primary and secondary) is an unsatisfactory measure of human capital. This may explain why the two measures are jointly significant in all columns but not individually, and also why they tend to have opposite signs in most columns. We investigated robustness of the results to a different measure of education, enrolment. Primary and secondary school gross enrolment rate is calculated as the number of students attending primary or secondary school as a proportion of the population in the typical age group for that level of schooling.[73] Enrolment was a popular measure particularly in early convergence studies (Mankiw et al. 1992, Levine and Renelt 1992, and Sala-i-Martin 1997). The OECD (2004:141) states that "[i]nformation on enrolment rates at various levels of education provides a picture of the structure of different education systems, as well as of access to educational opportunities in those systems." Enrolment measures investment in human capital whereas we need a measure of the current skills of the workforce however this is a useful robustness check. A practical advantage of the enrolment measure is also that data is available

[73] Data on the net enrolment rate which takes into account the actual age of students is not available for our sample.

for one additional country, Luxembourg.[74] The data is from Barro and Lee (1994) for 5-year intervals. We use the same method as for years of schooling to impute values for the intervening years. Primary enrolment varies very little in our dataset (it is never below 90 per cent)[75], but secondary enrolment increases considerably during the period e.g. for Australia from 29 per cent in 1960 to 74 per cent in 1985. Standard deviations of primary and secondary enrolment are 2 and 18 percentage points respectively. We might expect identification problems for a primary education effect (when taking first differences the value is very close to zero).

Primary enrolment is positive and significant when the dependent variable is $-\Delta$COH/COH or BOF share, and also BOF capacity but only using K as the time control. The two variables are jointly significant at 5 per cent and 10 per cent in the first two cases but not using BOF capacity. They are not significant in the equivalent of columns 1, 2 and 6. Otherwise the results are robust to the change of human capital measure. In particular, the results regarding the international effect are qualitatively the same as using schooling: extent of use elsewhere is negative and significant (at 10 per cent in the equivalent of column 2) except positive and significant in the equivalent of column 5. Also, the point estimates are well within the confidence intervals of the estimates in Table 5.2. The exception is ΔCBOF/COH for which only the confidence intervals overlap.[76] Other point estimates change somewhat but also these results are qualitatively similar to when years of schooling is used.

---

[74] The dummy for the highest observed values of ΔCBOF/COH now includes 1978 for Luxembourg as well as 1973 for Japan.
[75] Including Finland would introduce more variation but this is not possible because data is not available for the nominal interest rate measure that we use.
[76] The coefficient on A% using enrolment is below the lower boundary of the confidence interval using schooling.

We conclude that the main results are robust to an alternative measure of human capital and to the inclusion of data from an additional country, Luxembourg. Conceptually years of schooling is a better measure than enrolment and this is reflected in the weaker statistical significance of enrolment. However we also note that using enrolment primary education appears to be more important whereas for years of schooling secondary education producers the bigger estimated effects.

5.5.4 GDP

The coefficient on world GDP is positive in columns 1, 2, 5 and 6 which is consistent with the hypothesis that the level of demand for steel, proxied by world real GDP, is positively related to diffusion. The average annual change in (log) world GDP in our sample is 0.04 which corresponds to a growth rate of four per cent per annum. The estimated ceteris paribus effect of this increase is that ΔCBOF/COH increases by 37 percentage points (confidence interval (+18, +55)) in the short run and by 33 points (+16, +51) in the long run. The respective estimates in column 2 are 28 points (+11, +45) in the short run and 23 points (+9, +37) in the long run. The magnitude of the estimates is very large given that the average value of ΔCBOF/COH is just 18 per cent (although the standard deviation of ΔCBOF/COH is large at 1.6 or 160 per cent; see Table 5.1). A large effect is also suggested by results in column 6 although this estimate is very imprecise. The ceteris paribus effect of a 4 per cent growth in GDP is that BOF capacity increases by 1.2M tons (-93,000, 2.6M) in the short run and by 6.5M (+1.1M, +12M) in the long run. Compared to the sample average BOF tonnage of 10M, these estimates are very large. The column 5 estimates are quite different. The estimated increase in capacity is only 7 tons (+3.6, +10) in the short run and 48 tons (-8, +104) in the long run. Also, the results suggest that an increase in GDP growth has a significant negative effect on the rate -ΔCOH/COH which slows down by 11 percentage points (-19, -3.7) in the short run and by 15 points (-27, -2.9) in the long run. This is also a large effect (although imprecisely estimated).

The results can be explained if we consider that it is BOF capacity which responds to changes in demand whereas open hearth capacity mainly reflects the diffusion process. We note that GDP has a positive coefficient in the regressions which use either BOF capacity or change in BOF capacity (as a proportion of remaining open hearth capacity) as the dependent variable, whereas GDP does not have this effect if we examine the fall in open hearth capacity, or the share of BOF in total capacity. These two latter measures are perhaps more closely related to diffusion as a process of displacement. We consider that the GDP is picking up the effect of steel demand on overall capacity rather than the effect of demand on the displacement of open hearths by BOF.

A possible explanation for the negative coefficient in column 3 and the low precision of the other estimates is that (ln) world GDP is highly correlated with A%, A, and K. The correlation with A% is 0.98, with A 0.94, and with K -0.97. In general, multicollinearity may lead to individual coefficients not being identified which may then result in coefficient estimates having the wrong sign. In our case multicollinearity may then explain not only the negative coefficient on GDP in column 3 but also the positive coefficient on K in column 1 and lack of robustness of the BOF capacity estimates to the use of K (column 5). Indeed when we replace K and ΔK with time dummies, which reduces multicollinearity in the model, the coefficient estimates of A% or A and of GDP change. Unfortunately we have no other measure of the demand for steel available to us at this point.

In our view the results indicate that GDP is a successful proxy for steel demand in the sense that it has the correct sign and suggest an effect of significant magnitude in those regressions in which the dependent variable measures changes due to demand as well as the diffusion process per se. Similarly a lack of a positive effect in the other cases indicates that changes in demand affect overall capacity more than they affect the displacement of open hearths by BOF.

5.5.5 Real interest rate and the exchange rate

The nominal interest rate and inflation rate (CPI % change) are expected to jointly capture the effect of the real interest rate on the opportunity cost of adoption. There is evidence of such an effect in columns 4-6. The two measures are jointly significant (test statistic "Wald (r,CPI)") and both are also individually significant at 5 per cent in two columns. However the signs are not robust and only consistent with our expectations in column 6 in which the interest rate is not individually significant.

The nominal interest rate (government bond yield) is individually significant with the expected negative sign in column 5. On average, the interest rate is 8 per cent in our sample (minimum 2.4, maximum 20). If this increases by one per cent, estimates in column 5 suggest that a one percentage point increase in the nominal interest rate reduces BOF capacity by 480 tons (-700, -25) in the short run and by 3300 tons (-7,000, +500) in the long run. This suggests an effect of little economic significance.

The inflation rate has the expected positive sign in columns 4 and 6. The consumer price index increases on average by 6 per cent annually in our sample. If this rate increases to 7 per cent, the estimated effect is that BOF share increases only by 0.16 percentage points (-0.04, +0.36) in the short run and by 1.9 percentage points (-0.22, +4.1) in the long run. These estimates are imprecise and only the long-run point estimate suggests that inflation has an economically significant effect. BOF tonnage is estimated to increase by 330,000 tons (+180,000, +480,000) in the short run and by 1.7M tons (+0.59M, +2.8M) in the long run. These suggest an effect of considerable magnitude.

We also estimated a model in which the nominal rate and inflation are combined into a single regressor, the real interest rate. We find that estimated magnitude of the real interest rate effect is economically significant in the equivalent of columns 4 and 6. The coefficient on the real interest rate is negative and statistically significant in the equivalent of columns 4-6 with point estimates and standard errors -0.191 (0.102; p-value 0.063); -6.93 (1.68); -32,300 (8,020) respectively. In the second case (BOF capacity with K and ΔK) the magnitude is negligible[77] however the other estimates suggest that a one percentage point increase in the real interest rate reduces BOF share by 0.19 percentage points (-0.39, 0) in the short run and by 3.2 percentage points (-6.9, +0.49) in the long run; and BOF tonnage by 320,000 tons (-480,000, -170,000) in the short run and by 1.7M tons (-2.8M, -0.57M) in the long run. In the equivalent of columns 1-3 the real interest rate is not statistically significant. We do not report the full results because they follow closely the findings of Table 5.2. In particular, other estimates in columns 4-6 are not considerably affected and all point estimates are within the confidence intervals in Table 5.2. The fit of the model is slightly better in the equivalent of columns 5 and 6 (equation

[77] The estimated short-run reduction in BOF tonnage is 69 tons (-100, -36) and 470 tons (-1,000, +69) in the long run, given a one percentage point increase in the real interest rate.

standard errors 2899 and 2839 vs. 2904 and 2850 in Table 5.2). We conclude that there is some evidence of an interest rate effect which is of an economically significant magnitude, however the effect is not very precisely estimated.

There is no evidence of an exchange rate effect either in terms of the level or the change in the rate. Only one point estimate is significant, and even then only at 10 per cent (column 1). This has the expected negative sign but the magnitude of the effect is very small. There are several plausible explanations for why the variables are not statistically significant. Firstly, most of the variation in the exchange rate is due to a weak and volatile exchange rate in a few countries. Secondly, there is a theoretical explanation namely that the effect of the exchange rate on diffusion is unclear in our theoretical model. Recall that the exchange rate affects the optimal adoption date through two channels: the price of new technology (furnaces) and the price of output (steel). The impact through the second channel is unclear because exchange rates affect not only the domestic optimal adoption date but also producers elsewhere. Our hypothesis of a negative effect (that a weak exchange rate is related with a low extent of use) is based on the assumption that the exchange rate mainly affects the price of furnaces. However the interdependence that characterises the theoretical model suggests a complex relationship and the empirical evidence does not appear to shed any further light on this.

Because of the lack of statistical significance we experimented with a reduced model in which no exchange rate terms are included. However, excluding the exchange rate in this way appears to create omitted variable bias in particular on the estimate of the nominal interest rate.[78] This indicates to us that although there is no statistical evidence of an exchange rate there are also not sufficient reasons to exclude it from the model altogether.

5.6 Robustness checks

As a first robustness check, we estimated a basic model with a subset of explanatory variables. The model includes a lagged dependent variable, A or A%, time control(s) and a constant. This exercise shows that the international stock effect is not identified unless some other explanatory variables, suggested by theory, are included. The objective was to investigate how robust the international effect is to the exclusion of the eight additional regressors (world GDP, schooling, political instability, real interest rate, exchange rate). The sample size is larger without these regressors so that we can use the full data set from 1952 to 1985. The exercise also provides an opportunity to discuss the choice between time dummies and the nonlinear proxy K in some more detail. Estimates of the basic model are presented in Table 5.4. For each dependent variable, two sets of results are presented: one with K and ΔK, and one with biannual time dummies. In columns 7a and 7b a dummy is included for the two highest values of the dependent variable.

As in the general model, there is evidence that the domestic and foreign stock effects exist and that they are different from each other. The variables S(t-1) and A(t-1) or A%(t-1) are jointly and individually significant with different coefficients (see "Wald

---

[78] This is unlikely to be due to multicollinearity since the exchange rate is not highly correlated with any of the other regressors.

S,A" and "χ2 LinRes(S=A)" for test statistics) in all columns except 7a and 8b.[79] The lagged dependent variable is significant and positive in columns 8a-10b but negative in columns 7a and 7b which suggests the same conclusion as earlier namely that the domestic stock effect is not negative as was expected. The point estimates are within the confidence intervals in Table 5.2.

The sign of the international effect is not as robust as in the general model. There are only three estimates which are negative and significant, three are significantly positive, and two are not different from zero. This result reminds us of the lack of robustness of the estimated international effect in Chapter 3, where the international effect was positive in a third of the sample, negative in another third, and insignificant in the last third. The point estimates in columns 7b, 8a, 10a and 10b are within the confidence intervals in Table 5.2. There does not appear to be any correlation between the sign of the international effect and the choice of time controls.

More generally it is clear that the results of this basic model are unsatisfactory. The consistency of GMM is in doubt because there is significant second-order serial correlation in columns 9a and 9b, and the Sargan statistic is far above the critical values in all columns except 8a and 8b. The evidence for serial correlation may also indicate that the slope parameters vary across countries; Durlauf et al. (2005) suggest splitting the sample into groups that are more likely to share similar parameter values. We do not explore this further but the considerably longer time-series used in this exercise may explain why parameter homogeneity is suggested

---

[79] Note that the linear restriction test only makes sense when S and A are measured in the same units, i.e. columns 9a-10b.

here but not in the general model. Surprisingly the equation standard error is smaller in columns 10a and 10b than in columns 5 and 6 of Table 5.2; this is difficult to explain given that the sample here is considerably larger and there are eight less regressors, which are jointly significant in the general model. In our view this last finding is a further reason for some caution over the regression results in which BOF capacity is the dependent variable.

Table 5.4 Basic model results

Explanatory variables	Dependent variable (S)															
	ΔCBOF/COH				-ΔCOH/COH				BOF share				BOF capacity			
	7a		7b		8a		8b		9a		9b		10a		10b	
S (t-1)	-0.189	(0.139)	-0.277 *	(0.158)	0.326 **	(0.117)	0.276 *	(0.154)	0.94 **	(0.0434)	0.973 **	(0.0472)	0.933 **	(0.00406)	0.887 **	(0.0196)
A% (t-1)	-1.12	(0.758)	-9.37 **	(3.67)	-		-		0.112 **	(0.0551)	0.349 *	(0.205)	-		-	
A (t-1)	-		-		-0.00139 **	(4.68E-04)	0.00169	(0.00113)	-		-		0.00659 **	(0.00211)	-0.0416 **	(0.0186)
$K=2^{-0.1t}$	2.89 **	(1.44)	-		2.08 **	(0.582)	-		-0.259 **	(0.0551)	-		0.729 **	(0.322)	-	
ΔK	-0.207	(0.103)			-0.149 **	(0.0418)			0.0186 **	(0.00397)			-0.0523 **	(0.0231)		
time dummies	no		yes		no		yes		no		yes		no		yes	
constant	0.151 **	(0.0745)	-0.0144	(0.0406)	0.0753 **	(0.0182)	-0.525 **	(0.228)	-0.00792 **	(0.00123)	-0.00345	(0.00341)	-38.7 **	(17.1)	326	(229)
No. of obs.	349		349		261		261		462		462		462		462	
No. of countries	15		15		15		15		15		15		15		15	
AR(1)	-1.9 *		-2.7 **		-2.9 **		-2.5 **		-2.7 **		-2.9 **		-1.38		-1.4	
AR(2)	0.92		0.82		0.21		-0.55		-2.0 **		-2.3 **		-0.684		-0.58	
Sargan test	448 **	(226)	449 **	(226)	246	(228)	232	(228)	538 **	(234)	509 **	(234)	716 **	(234)	586 **	(234)
equation st.error	0.898		0.875		0.233		0.228		0.0416		0.0426		2230		2180	
Wald (S,A)	2.4	(2)	6.8 **	(2)	47 **	(2)	4.3	(2)	2,280 **	(2)	700 **	(2)	119,000 **	(2)	23,500 **	(2)
Wald (const+t)	4.1	(3)	1400 **	(14)	17 **	(3)	3,000 **	(15)	42 **	(3)	100 **	(13)	5.1	(3)	790 **	(14)
Wald (K,ΔK)	4	(2)			13 **	(2)			22 **	(2)			5.1 *	(2)		
χ^2 LinRes(S=A)	-		-		-		-		76 **	(1)	13 **	(1)	25,600 **	(1)	23,500 **	(1)

Notes: * significant at 10 per cent, ** significant at 5 per cent. Standard errors in brackets. A (t-1) is measured in 1M tons in columns 8a and 8b and in 1000 tons in columns 10a and 10b. A dummy for two observations is included in columns 7a and 7b (Japan 1973, Luxembourg 1978).

Based on the results of the general model (Table 5.2), the basic model estimated here, and the general model estimated using BOF growth rate as the dependent variable (see below), the nonlinear time control K tends to perform well but in some cases time dummies are better, particularly with BOF capacity. In Table 5.4 we report test statistics for the joint significance of the time controls and the constant ("Wald (const+t)"), and the two terms K and ΔK where appropriate (Wald (K,ΔK). The constant is included because it can be interpreted as a linear time trend in the levels model. Unsurprisingly the time dummies are significant in all columns. K, which we use a proxy for the level of adoption cost, is significant in all columns and so is ΔK except in 7a. However, the two terms are not jointly significant when the dependent variable is –ΔCOH/COH and only marginally significant using BOF capacity.

A robust finding is that the international effect is estimated with greater precision when K is used, i.e. in the "a"-columns in Table 5.4. The standard error is between 2.5 to nearly 9 times as large in the time dummy regressions. K appears to be a good choice over time dummies even when it has a positive coefficient estimate. Why this is so is not immediately clear, as a positive coefficient contradicts the interpretation that K acts as a proxy for the falling price of furnaces over time. The equation standard error cannot be used to distinguish between the K and time dummies because it increases with the number of explanatory variables. With BOF share as the dependent variable, the equation standard error is higher with time dummies which is a strong suggestion that K is a good time control in this case. Finally, we note that the consistency of GMM is not sensitive to the choice of time control. We conclude that although time dummies work well in column 6, generally we are confident in our use of K as the time control.

The largest difference in the precision of estimates is in columns 10a and 10b. The weakness of results in 10a reflects what we found in the general model namely that estimates in column 5 are repeatedly different from those in other columns. It seems clear that the use of K is not a good choice when the dependent variable is BOF capacity. There is possibly an issue with the estimation method in this case because the Sargan statistic, which rejects the overidentifying restrictions in all cases using this dependent variable, is especially large in column 10a (and relatively high in column 5). We suggests that rather than conclude that time dummies are preferred the issue is more generally why the results are so sensitive to the choice of time control terms. For this purpose we estimated the general model (i.e. equivalent of column 5) without the term ΔK and found that the results are considerably altered. They are closer to those obtained from other columns and in particular we find that the coefficient on A is negative and significant although only at 10 per cent. The magnitude of the other estimated effects such as world GDP is however still very small compared to the other columns.

In other columns, a likely explanation for the difference in standard errors for A or A% is that time dummies capture much of the time-series variation in S_{it} and the amount left to be explained by A or A% may be small. In Table 5.4, this is apparent in how the Wald test statistic for the joint significance of S and A or A% is smaller in the “b”-columns than the “a”-columns except in 7a and 7b. Because of the way they are constructed (world diffusion less country i’s contribution), A% in particular but also A vary relatively little across countries. This means that they must explain enough of the time-series variation in S_{it} in order to be statistically significant. Because K imposes structure on the time effect whereas time dummies do not, K leaves more of the time-series variation unexplained which might explain why A and A% are estimated with higher precision.

It appears to us that the relationship between time, the foreign and domestic stock effects and domestic diffusion is somewhat different in the case of the "relative rate of growth", ΔCBOF/COH, than the other diffusion measures. The main difference is of course that the domestic effect has a negative sign. We were expecting a negative coefficient in all regressions because in the theoretical model both domestic and foreign use have the same stock effect. Why a positive coefficient is obtained in other cases is discussed in section 5.7 below. However, it is also worth pointing out that this measure of diffusion takes on potentially very high values such as Japan 1973 and Luxembourg 1978 in our sample (the latter is not used in the general model) if the last open hearth tonnages that are reported are very small. Indeed we find that the coefficient estimate for A% is sensitive to the inclusion of the dummy without which the coefficient on A% is positive (+1.12) and the equation standard error is very high (0.785) in the equivalent of column 2 in the general model.

We conducted two further checks of the robustness of the general model: first, using the growth rate of BOF as the dependent variable; and second, we estimated the model over the earlier period 1952-1973 which is possible by using a different education measure. The full results are not reported but we outline the findings here.

In the main analysis, we have not included results which use the growth rate of basic oxygen capacity as the dependent variable. The reason is that we are concerned that these results are particularly sensitive to any weaknesses that our proxy, world GDP, suffers from because the growth rate of BOF tonnage is expected to be particularly sensitive to annual fluctuation in the demand for steel. We estimated the general model and the results support our main argument that there is a statistically significant negative international effect. In this regression extent of use elsewhere is measured by BOF proportion (A%). The coefficient estimate (and standard error) is -0.953 (0.388) using K and ΔK as the time controls. The estimated effect is economically significant relative to the average annual BOF growth rate, 26 per cent in our sample: a three percentage point increase in A% is estimated to reduce BOF growth by 2.9 percentage points (confidence interval (-5.1, -0.58) points). The lagged dependent variable is not significant (point estimate 0.036 and standard error 0.027) although jointly the two lagged variables are significant. This finding of a strong international effect and relatively weak domestic effect is similar to the results of columns 1 and 2 where the dependent variable has the same numerator as here.

The eight regressors of the general model are jointly significant, as are the nominal interest rate and CPI percentage changes, however the schooling variables and political instability are not jointly significant. Regressors which are individually statistically significant are: world real GDP (+), nominal interest rate (-), inflation (+), and K (+) and ΔK (-). Since the coefficient on K is positive we also ran the regression using biannual time dummies. Here too the estimate of A% is very imprecise (standard error is 1.6 but only 0.39 using K) and in this case the consequence of using time dummies is that the coefficient on A% is not significantly different from zero. The eight other explanatory variables are jointly significant but only world GDP is individually significant with a positive sign. In the basic model without the eight variables A% is negative and significant (point estimate -0.517, standard error 0.273) and the lagged dependent is positive and significant but only if we use K and ΔK; using time dummies both lagged variables are positive.

The case of BOF growth rate suggests that a negative international effect is identified in a model which includes additional regressors suggested by theory and a nonlinear time control rather than time dummies. There is also evidence of a significant world GDP effect and a real interest rate effect but no evidence that schooling or the other variables are significant.

Finally, an alternative enrolment measure provided in HCCTAD from the Banks (1976) dataset offered an opportunity to estimate the general model for the period 1952-1973. Our objective was to test whether our results regarding the international stock effect hold when: i) data from the first decade of diffusion is included; and ii) we only use data up to the recession which began in 1975. Unfortunately, data on political instability is not available before 1960 so the estimates are potentially exposed to omitted variable bias. We ran the regressions corresponding to Table 5.2 (without political instability) using the alternative primary and secondary enrolment measures in place of the years of schooling.[80]

There is evidence of a negative international effect only in two regressions, corresponding to columns 1 and 6. Using ΔCBOF/COH as the dependent variable, the point estimate is -1.34 (standard error 0.451) and using BOF capacity the estimate is -0.061 (0.0317, p-value 0.056). These are within the confidence intervals

[80] The enrolment measure which is available before 1960 is not compatible with the measure which begins in 1960.

in Table 5.2. The lagged dependent variable is not significant in the former case but in the latter the coefficient is close to unity and therefore the long-run coefficient on A is very large, -1.85, which is nine times the size of the estimate in column 6 (see Table 5.3). At least one of the enrolment measures is statistically significant in three of the columns (using BOF capacity and ΔCBOF/COH).

The evidence for an international effect is weak in this first decade of international diffusion (the 1950s) since evidence is only found for two of the four diffusion measures. This does not necessarily suggest that the effect is stronger later – indeed one of the estimates suggests a considerably larger effect than in the later period – because the exercise cannot be considered a reliable robustness check. The estimates are exposed to omitted variable bias because political instability is excluded,[81] and the sample size is very small (only 104 for -ΔCOH/COH). This is because the time series are short, especially for the two change measures (the shortest time-series is only three years for ΔCBOF/COH and -ΔCOH/COH). Finally, there is a further reason for caution because as discussed earlier the literature suggests that 1960 can be considered a threshold after which the superiority of BOF technology over the open hearth has been firmly established.

5.7 Conclusion

We have estimated a model of the diffusion of the basic oxygen furnace in a sample of OECD countries. Our empirical model is based on the decision-theoretic model

[81] We found some evidence of such bias in the exchange rate. In this exercise, the exchange rate is statistically significant in the equivalent of columns 3-7 unlike in any of the regressions reported earlier (although the sign is not robust across diffusion measures). This statistical significance disappears however if we estimate the model using enrolment and political instability for the period 1960-1974. This may indicate that the significance of the exchange rate is due to bias from omitting political instability.

developed in the previous chapter, from which we draw a number of hypotheses regarding the country-level determinants of diffusion. The main objective is to test the hypothesis that there exists a negative international stock effect. The argument is that an increase in the extent to which the BOF is used by producers abroad changes the arbitrage conditions of domestic producers so that postponing adoption becomes more attractive, ceteris paribus. The empirical evidence strongly supports such a relationship, and our estimates suggest that the magnitude of the international effect is also economically significant. This conclusion is based on examining four different measures of domestic diffusion, two of which are measures of the level of diffusion and two which measure change in the extent of use.

We also find that the relationship between past and present domestic diffusion is positive for three of the four diffusion measures. This contradicts the hypothesis that the stock effect depends on the total extent of use in the world. In the theoretical model, the benefit of using BOF as opposed to the open hearth declines with each additional adoption whether at home or abroad. Thus we would expect there to be a negative domestic stock effect as well as an international one. However the lagged dependent variable is in fact close to unity for the two level measures of diffusion, and also significantly positive for one of the measures of change. While the level measures can be considered naturally persistent over time and this makes the finding unsurprising, the positive coefficients cannot be explained by our theoretical model.

It seems to us that this finding regarding the lagged dependent variable should be interpreted in light of previous empirical studies of intra-country diffusion. Our short review of empirical studies in section 1.3.2 highlighted the lack of previous evidence for a stock effect. This is usually accounted for by the fact that a single measure, the current extent of use, is used to capture not only stock but also epidemic and order effects. The epidemic modelling tradition argues that the expected relationship between current and future (domestic) use is positive because of information-spreading (see Chapter 3). In most studies the empirical evidence supports this relationship, and indeed our results are similar in this respect. Also in so-called spillover studies the argument is made that informational or learning spillovers explain why research or use of a technology by one firm has a positive effect on, for example, the productivity of another firm. An interesting example is Branstetter (2001), who studies information spillovers from research by other firms and expects to find positive relationships both domestically and internationally. In his study only the domestic effect is statistically significant, with the expected positive sign. He suggests that this may be because the positive spillovers are overcome by the effects of international rivalry. Our hypothesis is that both domestic and international diffusion has a negative effect because of competition in the international output market. We find evidence that both effects are significant but that the domestic effect is positive. Although unexplained by our theoretical model (or the model of Reinganum 1981b) in light of the literature the finding is unsurprising and suggests that the finding may be explained by expanding our theoretical model to include some degree of uncertainty.

We included a number of additional regressors to capture cross-country differences in the optimal adoption date. The results support the hypothesis that the level of human capital, here measured by years of schooling, affects diffusion. The sign of the effect cannot be established nor a conclusion made about which of primary and secondary education matters more. The two measures are jointly significant however for three of the four diffusion measures. There is also some evidence of the importance of political instability, measured here by the number of assassinations and revolutions (or attempts at either) however we are somewhat cautious about interpreting this finding. Variation in this measure is limited to a few countries mainly in the 1970s and for this reason statistical significance (with a sign that is not robust) may reflect some other factor common to those countries and time periods which we do not observe. Although we cannot establish their individual effects in this data the joint significance the education measures and political instability is evidence of their importance for diffusion. This finding is interesting given that such measures of institutional differences have been largely ignored in the literature.

There is also evidence that the domestic real interest rate is a significant determinant of the extent of use. The nominal interest rate and inflation, which are measured by the government bond yield and percentage change in the consumer price index respectively, are jointly significant when the dependent variable is BOF proportion or BOF capacity. The signs of these variables are not robust. However, in a regression where the two are combined into a single real interest rate measure, the variable has the expected negative sign and is statistically significant. This is consistent with the hypothesis that a high opportunity cost of adoption is an incentive to postpone adoption. There is no evidence that the domestic exchange rate a significant effect on diffusion however we have argued that in light of the complex role the exchange rate plays in the theoretical model this is not very surprising.

We find that estimates of log real world GDP are relatively robust across different models. We use world GDP as a proxy for the demand for steel because no direct measure is available and it is computed simply as the sum of real GDP in the 15 countries in our sample. The estimated effect is positive and significant in those regressions in which the dependent variable can be considered most likely to reflect not only the process of diffusion – the displacement of open hearths – but also changes in total capacity as a response to changes in demand. We include real GDP in order to control for the latter type of changes in BOF tonnage. The same positive relationship between GDP and diffusion is not found with measures that mostly reflect diffusion as the displacement of old technology, namely the negative growth rate of open hearth tonnage and the BOF proportion. We interpret this as an indication that changes in demand affect total production capacity rather than the diffusion process as such.

Accounting for a time trend in intra-country diffusion is necessary to identify the international effect, and we find that an arbitrary nonlinear term $K(t)=2^{-0.1t}$ performs better than (biannual) time dummies. In particular, the estimates of the international effect are considerably more accurate, which we explain by that time dummies leave relatively little time-series variation to be explained by other variables. However the coefficient on K is not negative as expected given that it is intended to capture the falling price of adoption over time. The change in K has the opposite sign as expected and is significant in those regressions in which K is also significant. The results suggest to us that the use of a nonlinear control even if arbitrary is preferred because it imposes structure on the hypothesised time effect.

A limitation of our study is the focus on the demand for technology, which is consistent with the bias in the literature to treat the supply-side very simplistically. It would be interesting to investigate whether our finding that the past domestic extent of use has a positive effect can be explained by the existence of some local (i.e. common to all producers within a country) adoption costs that are, to some extent, endogenous to the extent of domestic diffusion. By this we mean that the assumption of an exogenous world adoption cost is maintained but an element of total costs is endogenous locally. Possible reasons would include scale or learning effects in the supply of new technology (e.g. transport costs), and learning by adopters.

A natural extension of the work here would be to build a more detailed model of the specific case of the BOF with clear hypotheses about the informational or epidemic effects relevant to this diffusion process. Thus a way forward would be to develop hypothesis about how such effects may have been important in the case of BOF diffusion. This approach is likely to introduce technology-specific explanatory variables into the empirical model. There has been a tendency in some of the empirical literature, in particular policy-oriented studies and marketing literature, to include a number (sometimes large) of technology-specific explanatory variables in addition to measures of rank, stock, order and epidemic effects (if any). The objective of this approach is to increase the power of the model so that predictions can be made about future diffusion and the effects of any policy change. An example of such a study is Hollenstein and Woerter (2008:554), who argue that epidemic, rank, stock and order effect "are too general to fully capture technology-specific benefits".

In the present study we have purposefully kept the empirical models (in this Chapter and in Chapter 3) general so that the link with theoretical hypotheses, formulated on a general level not specifically for the present technology, is as clear and direct as possible. We have not attempted to obtain best possible "fit" to the data but have rather valued parsimony and the chance to test one particular theoretical hypothesis, the international stock effect. The weakness of technology-specific regressors is that their theoretical connection is often weak (or not explicit), and that these variables typically end up explaining most of the variation in the dependent variable. This reduces the relevance of such a study to a reader not interested in that particular technology and also the potential of the study to contribute to theoretical discussion; although some studies have successfully linked technology-specific variables to theory (e.g. Colombo and Mosconi, 1995). Also, even a degree of detail far beyond the data we have available, does not guarantee that the model is "satisfactory". An example is the difficulty in accounting for the effects of changes in steel demand on the diffusion of the BOF. Beeson and Giarratani (1998) construct a plant-specific demand measure that is an index of end product prices, the plant's product mix, and the distance from the plant to the market. Even this amount of detail is, according to the authors, insufficient to capture variation in true demand for the plant's output in the United States during the 1970s and 1980s.

Finally, an alternative to panel data methods is separate time-series regressions for each country as in Chapter 3. The advantage of panel methods over separate regression is that information contained in cross-country variation can be used. Parameter estimates are likely to be more efficient. It is worth noting that time-series as long as what we had available for steam- and motor ships are very rare in the diffusion literature, and the BOF series are also relatively long. A panel model is restrictive in that it imposes parameter homogeneity across units, however the serial correlation tests do not suggest that there is any evidence that the assumption of homogeneity is violated (see 5.5.1).

6 Conclusion

6.1 An overview

It is only through the diffusion process that the benefits of technological progress are fully realised. This study was motivated in part by the relatively limited emphasis placed upon inter-country diffusion in the economics literature on diffusion as opposed to the much more studied process of intra-country diffusion. We began the study in Chapter 2 by devising a measure of diffusion that can be decomposed to the intensive (intra-country) and extensive (inter-country) margins. Using a number of historical examples we empirically demonstrated the relative importance of the two margins over the whole diffusion process. These findings lead one immediately to consider how the extensive and intensive margins are related.

An epidemic model of diffusion was then developed in which the relationship between international and domestic use of a technology is much the same as in the conditional convergence literature in macroeconomics (see below): diffusion in other countries has a positive effect on diffusion at home because there is some positive knowledge spillover effect. So the greater the extent of diffusion in the world, the greater the amount of knowledge available about the new technology and thus the lesser the degree of uncertainty about it, which encourages further adoption. In Chapter 3 we tested this hypothesis in the context of steam- and motor ship diffusion (using data from the Historical Cross-Country Technology Adoption Dataset or HCCTAD). The empirical evidence is suggestive of such an international effect, but the results are inconclusive because the effect has the expected positive sign only in a third of our sample of countries. In another third the relationship between international and domestic diffusion is also statistically significant but negative. This result is inconsistent with the theoretical notion of information spillovers if we maintain that steam- and motor ships are indeed the superior technology. In the epidemic framework inter-country diffusion can only have a positive effect on further use because information is inherently encouraging of adoption. In the context of the particular technology that we studied, the negative relationship is possibly explained by the nature of the relative benefit of steam- and motor ships over sailing ships which (previous studies suggest) was not uniform across markets.

The theoretical prediction of the decision-theoretic model developed in Chapter 4 is quite opposite to that of the epidemic model. In decision-theoretic models such as this, it is misleading to talk of spillovers in the sense that the effect of adoption by others over time is internalised so it becomes an anticipated or expected part of the adoption timing problem faced by an individual (potential) adopter. In our particular model which builds upon Reinganum (1981b), the international effect is established by modelling how interaction between producers in the output market determines the effect of the diffusion of a process technology on the individual producer's adoption decision. We show that adoption by others discourages further adoptions through the so-called stock effect. The study contributes to the literature by: i) considering that the "stock effect" may have an international dimension as opposed to a purely domestic dimension; and ii) by considering the effects of producer heterogeneity on the determinants of international diffusion both at the intensive and extensive margins.

The empirical evidence for an international stock effect is stronger than the evidence of a positive effect in the epidemic model. In Chapter 5 we studied the diffusion of the basic oxygen furnace in a panel of countries and found that the international effect is indeed negative. The result appears to be robust to a number of different measures of diffusion. However, we also find that there is a positive relationship between domestic diffusion and past domestic extent of use which is inconsistent with the theoretical model in which the location of diffusion does not matter. Our results also open up new questions for further research as this finding suggests that the model can be improved by considering the geographical dimension of diffusion, for which there exists a growing body of literature.

The model developed in Chapters 4 and 5 is also a framework for the analysis of structural and policy variables in international diffusion. We use the arbitrage condition to argue that the optimal adoption date depends on the characteristics of the technology and the market, producer heterogeneity, and also country factors which affect production and adoption costs more generally. In the empirical study we examine eight such variables, one (world GDP) which is common to producers in all countries, and seven others which are hypothesised to affect the costs and benefits of postponing adoption within a specific country. Although only some of the individual effects are identified (world GDP and real interest rate) the evidence supports the hypothesis that the eight variables are jointly significant determinants of intra-country diffusion. In particular, we find that human capital (years of primary and secondary schooling) and political instability are jointly significant. This suggests that institutional factors have a greater role to play in diffusion than that assigned to them in the literature. There is a tendency in the empirical literature to examine single-country diffusion processes and also technologies which are relatively fast to diffuse (i.e. the time-series are short) which goes some way to explain the relative lack of interest in social and political variables which by their nature vary little over time. A panel data model however offers the opportunity to identify institutional effects by utilising both cross-sectional and time-series variation.

The Cournot model of Chapter 4 is particularly interesting because it tells us that policy or institutional changes in one country will affect the use of new technology in that country but also in other countries as adopters abroad react to changes in the domestic adopters' environment. That is, factors which affect domestic diffusion are also potentially important for international diffusion. We discussed at length how the exchange rate impacts not only upon the domestic cost and profitability of adoption but when both output and technology are traded in a world market the exchange rates of all other countries also matter through the interdependence of producers' decision-making. This example underlines the importance of considering the international dimension of diffusion even when the aim of research may be explaining the diffusion process in a particular country.

This study contributes to both the theoretical and empirical diffusion literatures by proposing ways in which the international dimension of diffusion can be incorporated into the analysis to an extent which has not been done before. The theoretical aim of this study has not been to propose a novel model that encompasses the existing theoretical approaches. Nor has the empirical aim been to present models of particular diffusion processes that would explain those processes better than any previous studies. What we have done is applied three well-known models of diffusion to an international context, derived hypotheses about the link between international and domestic diffusion, and presented empirical evidence of that link. We have demonstrated that the theoretical insights provided by such extended models applied to the international setting are interesting and considerable. Thus the study demonstrates the importance of considering the international dimension of diffusion and we suggest concrete ways in which this can be done both theoretically and empirically.

6.2 Further linkages

In addition to our diffusion-related findings, it is informative to reflect on the importance of our findings in relation to the literature on economic growth and convergence. In macroeconomic models of economic growth, technological change takes centre stage. In the neoclassical Solow-Swan growth model countries which share the same structural characteristics converge to a common steady state in the long-run. The steady state is characterised by a rate of per capita GDP growth which depends exclusively on the rate of technological progress. The prediction of the model that has been the focus of many empirical studies is that countries which initially start with a low level of per capita output will experience high rates of growth on the path towards the steady state. This catch-up is enabled by the transfer of technological knowledge which is assumed to be a public good. A classic study of the convergence hypothesis is Baumol (1986) who regresses change in per capita GDP between 1870 and 1979 on the initial value and finds evidence of convergence among industrialised market economies.

Empirically, correlation between growth and initial income levels is weak and the evidence suggests divergence rather than convergence of per capita income levels in the world as a whole (e.g. Mankiw et al. 1992). Baumol argued that there are distinct "convergence clubs", each of which is characterised by the same underlying structural parameters. In other words, convergence is conditional on the steady state. De Long (1988) showed that measurement error in the 1870 figures eliminates most of the convergence that Baumol found. He also brought into the discussion Abramovitz' (1986) notion of social capability and argued that convergence clubs are formed on the basis of a country's ability to assimilate or grasp the benefits of new technology. Patent laws, intellectual property rights, and other aspects of the social infrastructure are then key determinants of a country's steady state level of per capita income.

The treatment of technological change in the neoclassical growth model is unsatisfactory in many respects. Most obviously, long-run growth depends on the rate of technological progress which is simply a parameter. Also, the key prediction of convergence depends on the existence of technological spillovers. Technology transfer across countries is assumed to happen "automatically" because technological knowledge is a public good. A large body of literature has developed in which the production of new knowledge and its transfer across borders is explicitly modelled. In theoretical models such as Grossman and Helpman's (1993) North-South model, the role of patent rights is highlighted as an incentive to innovate and trade is presented as a channel of technology transfer. In the empirical literature, numerous studies of technological spillovers can be found which typically examine R&D data and productivity data, with mixed results. Branstetter (2001) for example finds no evidence that firms in one country benefit from research in another country, although he does show that there are intra-country spillovers. The economic growth literature therefore assigns great importance to the modelling and empirical study of innovation on the one hand, and the extensive margin of international diffusion on the other hand, but rarely study this per se.

The models of diffusion developed in this thesis suggest that there are “spillovers” in the sense that international diffusion affects domestic diffusion. This can be seen to be the first necessary condition for convergence that is however not made explicit in models of economic growth. Inter-country diffusion of technology is not sufficient for convergence.

6.3 Boundaries of this study

There is a vast literature on technological diffusion which spans several disciplines. In this thesis we have drawn on the information-spreading and decision-theoretic approaches and thus purposefully left outside the discussion arguments and analyses for which considerable bodies of literature exist. We view these literatures as complementary to our analysis in the sense that they present avenues along which the work here can be extended, refined and further developed in the future. We now outline some of these avenues which are most interesting for the present study: supply-side modelling, path dependence and spatial diffusion models.

Together with most of the diffusion literature, this study focuses on the demand for new technology as opposed to supply. That is, our analysis concerns primarily the use of a technology not its production. Stoneman and Ireland (1983) developed an early model of the supply-side but few studies since (either empirical or theoretical) have developed this further. Much of the research on the supply of new technology is found outside the diffusion literature; for example, in macroeconomic models of technology transfer. Stoneman and Battisti (forthcoming) present an analytical framework in which supply and demand are treated as equally important. Their approach also brings together different levels of analysis, from intra-firm to the international level of diffusion, so these can be analysed within the same framework.

Taking a supply-side view implies in the first instance that the price and quality of new technology are endogenous to diffusion and determined through interaction between supply and demand (Stoneman and Ireland 1983, Ireland and Stoneman 1986). This contrasts with the assumption of exogenous technology price that we make in Chapters 4 and 5 and the model in Chapter 3 where suppliers act as an exogenous source of information. To model supply would offer a more complex view of the diffusion process and in the decision-theoretic model we have already discussed that there is scope for additional linkages between international and domestic diffusion through a modelling of the supply-side.

Two large bodies of literature which we do not consider in this study are path dependence and spatial diffusion models. It is quite clear that in both of the empirical examples of diffusion that we studied in some depth (Chapters 3 and 5) analytical methods from these literatures can be used to further refine the analysis. Path dependence draws attention to the limits that a firm's previous choices put on current decision-making. Paul David is one of leading authors (e.g. David 1985). It is argued that the set of technological options that a firm has at a given point in time depends on its history such as when, how and what technologies it has adopted in the past. Adopting technologies which require big changes to the organisation of production is considerably different than those that do not require such changes, and therefore past experience of such changes affects adoption costs. In general the path dependence approach calls for a model in which producers are heterogeneous in terms of adoption costs. The first step for us would be to take into account vintage capital, that is, the age of the current stock, which determines the costs of scrapping that base when new (process) technology is to be adopted. Then we could study more in depth the constraints that potential adopters, that is, ship owners and steelmakers in our case, faced in making the transition from the old to the new technology. One of the likely benefits of this approach is the development of a more detailed understanding of exactly which elements of the social and political environment are relevant for the adoption decision in the case of a particular technology.

A spatial view of diffusion suggests in the first instance that any measure of international diffusion should allow countries to have different weights depending on their “closeness” (geographical or more widely defined). Durlauf et al. (2005) writing in the growth economics context point out that there is no natural cross-sectional ordering and that the socioeconomic space in which distances should be measured is presumably determined by multiple channels. This applies to diffusion studies as well. In Chapter 3 we argued that our model could be extended by taking into account market segmentation. Differences across countries in terms of the markets that their fleet operates in could be used to assign weights to the contribution of each country to the international diffusion measure. Baptista (2000) for example builds a model where epidemic effects are geographically restricted (see also Baptista 1999 for a review). In Chapter 5 we indicated that there is likely to be cross-country correlation in the error term due to common effects unobserved by us or common shocks. The first step here would be to establish, from the literature, how the spatial correlation structure should be specified or estimated. It seems likely to us that sub-groups of countries may be defined by geographical proximity but again this approach would also give an opportunity to study the degree to which the market for steel is localized e.g. by allowing local demand shocks to affect countries differently. Together with supply-side modelling and path dependence, spatial models represents an interesting avenue for future studies.

7 References

Abramovitz, M. (1986) Catching Up, Forging Ahead, and Falling Behind. *The Journal of Economic History*, 46 (2): 385-406.

Acemoglu, D. and Robinson, J. A. (2001) A Theory of Political Transitions. *American Economic Review*, 91 (4): 938-963.

Adams, W. and Dirlam, J. B. (1966) Big Steel, Invention, and Innovation. *The Quarterly Journal of Economics*, 80 (2): 167-189.

Alderton, T. and Winchester, N. (2002) Globalisation and de-regulation in the maritime industry. *Marine Policy*, 26 (1): 35-43.

Alesina, A. and Rodrik, D. (1994) Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109 (2): 465-490.

Arellano, M. and Bond, S. (1991) Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* 58 (2): 277-297.

Banks, A. S. (1976) *Cross-National Time Series Data Archive*. Center for Social Analysis, State University of New York at Binghampton.

Banks, A. S. (1979) *Cross-National Time Series Data Archive*. Center for Social Analysis, State University of New York at Binghampton, September 1979, updated.

Baptista, R. (1999) The Diffusion of Process Innovations: A Selective Review. *International Journal of the Economics of Business*, 6 (1): 107-129.

Baptista, R. (2000) Do innovations diffuse faster within geographical clusters? *International Journal of Industrial Organization*, 18 (3): 515-535.

Barro, R. J. and Lee, J-W. (1994) *Barro-Lee Data Set for a Panel of 138 Countries*. Revised January 1994. http://www.nber.org/pub/barro.lee, accessed 3/9/2009.

Barro, R. J. and Lee, J-W. (2000) *International Data on Educational Attainment: Updates and Implications*. Manuscript, Harvard University, February 2000. http://www.cid.harvard.edu/ciddata/ciddata.html, accessed 3/9/2009.

Barro, R. J. and Sala-i-Martin, X. (1997) Technological diffusion, convergence and growth. *Journal of Economic Growth*, 2 (1): 1-27.

Barton, J. R. (1999) Flags of Convenience : Geoeconomics and Regulatory Minimisation. *Tijdschrift voor Economische en Sociale Geografie*, 90 (2): 142-155.

Bass, F. M. (1969) A New Product Growth for Model Consumer Durables. *Management Science*, 15 (5): 215-227.

Battisti, G. and Stoneman, P. (2003) Inter- and intra-firm effects in the diffusion of new process technology. *Research Policy*, 32 (9): 1641-1655.

Baumol, W. J. (1986) Productivity growth, convergence and welfare: What the long-run data show. *American Economic Review*, 76 (5): 1072-1085.

Beeson, P. and Giarratani, F. (1998) Spatial Aspects of Capacity Change by U.S. Integrated Steel Producers. *Journal of Regional Science*, 38 (3): 425-444.

Branstetter, L. G. (2001) Are knowledge spillovers international or intranational in scope? - Microeconometric evidence from the U.S. and Japan. *Journal of International Economics*, 53 (1): 53-79.

Breshanan, T. F. and Trajtenberg, M. (1995) General purpose technologies: "Engines of growth"? *Journal of Econometrics*, 65 83-108.

Broeze, F. J. A. (1975) The Cost of Distance: Shipping and the Early Australian Economy, 1788-1850. *The Economic History Review*, 28 (4): 582-597.

Canepa, A. and Stoneman, P. (2004) Comparative international diffusion: Patterns, determinants and policies. *Economics of Innovation and New Technology*, 13 (3): 279-298.

Caselli, F. and Coleman II, W. J. (2001) Cross-country technology diffusion: The case of computers. *American Economic Review*, 91 (2): 328-335.

Cochrane, J. H. (1996) A Cross-Sectional Test of an Investment-Based Asset Pricing Model. *Journal of Political Economy*, 104 (3): 572-621.

Cohn, R. L. (2005) The Transition from Sail to Steam in Immigration to the United States. *Journal of Economic History*, 65 (2): 469-495.

Colombo, M. G. and Mosconi, R. (1995) Complementarity and Cumulative Learning Effects in the Early Diffusion of Multiple Technologies. *Journal of Industrial Economics*, 43 (1): 13-48.

Comin, D. and Hobijn, B. (2003) *Historical Cross-Country Technology Adoption (HCCTA) Dataset*. http://www.nber.org/hccta, accessed 26/9/2009.

Comin, D. and Hobijn, B. (2004a) Cross-country technology adoption: making the theories face the facts. *Journal of Monetary Economics*, 51 39-83.

Comin, D. and Hobijn, B. (2004b) Neoclassical growth and the adoption of technologies. *NBER working paper series* No. 10733. Cambridge, MA: National Bureau of Economic Research, August 2004.

Comin, D. and Hobijn, B. (2006) An exploration of technology diffusion. *NBER working paper series* No. 12314. Cambridge, MA: National Bureau of Economic Research, June 2006.

Comin, D., Hobijn, B. and Rovito, E. (2006) Five facts you need to know about technology diffusion. *NBER working paper series* No. 11928. Cambridge, MA: National Bureau of Economic Research, January 2006.

Costa-Font, J. and Mossialos, E. (2005) Is dread of Genetically Modified Food associated with the consumers' demand for information? *Applied Economics Letters*, 12 (14): 859-863.

Crafts, N. (2004) Steam as a general purpose technology: a growth accounting perspective. *The Economic Journal*, 114 (495): 338-351.

David, P. A. (1985) Clio and the Economics of QWERTY. *American Economic Review*, 75 (2): 332-337.

David, P. A. and Olsen, T. E. (1986) Equilibrium dynamics of diffusion when incremental technological innovations are foreseen. *Ricerche Economiche*, 40 738-770.

Davies, S. (1979) *The diffusion of process innovations*. Cambridge etc.: Cambridge University Press.

De Long, J. B. (1988) Productivity growth, convergence and welfare: Comment. *American Economic Review*, 78 (5): 1138-1154.

Dekimpe, M. G., Parker, P. M. and Sarvary, M. (2000) Global diffusion of technological innovations: a coupled-hazard approach. *Journal of Marketing Research*, 37 (1): 47-59.

Dickey, D. A. and Fuller, W. A. (1979) Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74 (366): 427-431.

Dickey, D. A. and Fuller, W. A. (1981) Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, 49 (4): 1057-1072.

Dixit, A. K. and Pindyck, R. S. (1994) *Investment Under Uncertainty*. Princeton and Chichester: Princeton University Press.

Durlauf, S. N., Johnson, P. A. and Temple, J. R. W. (2005) Growth Econometrics. In: Aghion, P. and Durlauf, S. N. eds. *Handbook of Economic Growth*. Amsterdam: Elsevier.

Dutta, P. K., Lach, S. and Rustichini, A. (1995) Better late than early: Vertical differentiation in teh adoption of new technology. *Journal of Economics and Mangement Strategy*, 4 (4): 563-589.

Fudenberg, D. and Tirole, J. (1985) Preemption and Rent Equalization in the Adoption of New Technology. *Review of Economic Studies*, 52 (3): 383-401.

Fusaro, M. A. (2009) The rank, stock, order and epidemic effects of technology adoption: an empirical study of bounce protection programs. *Journal of Technoloy Transfer*, 34 (1): 24-42.

Geroski, P. (2000) Models of technology diffusion. *Research Policy*, 29 (4): 603-625.

Götz, G. (1999) Monopolistic Competition and the Diffusion of New Technology. *RAND Journal of Economics*, 30 (4): 679-693.

Graham, G. S. (1956) The Ascendancy of the Sailing Ship 1850-85. *The Economic History Review*, 9 (1): 74-88.

Granger, C. W. J. and Newbold, P. (1974) Spurious Regressions in Econometrics. *Journal of Econometrics*, 2 (2): 111-120.

Greenhill, B. (1968) *The Merchant Schooners*. Newton Abbot: David & Charles.

Griliches, Z. (1957) Hybird corn: An exploration in the economics of technological change. *Econometrica*, 25 (4): 501-522.

Grossman, G. M. and Helpman, E. (1993) *Innovation and growth in the global economy*. paper. Cambridge, Mass.: MIT Press.

Gruber, H. and Verboven, F. (2001) The diffusion of mobile telecommunications services in the European Union. *European Economic Review*, 45(3): 577-588.

Hall, B. H. (2004) Innovation and Diffusion. In: Fagerberg, J., Mowery, D. C. and Nelson, R. R. eds. *The Oxford Handbook of Innovation*. Oxford and New York: Oxford University Press.

Hannan, T. H. and McDowell, J. M. (1987) Rival Precedence and the Dynamics of Technology Adoption: An Empirical Analysis. *Economica*, 54 (214): 155-171.

Harley, C. K. (1971) The shift from sailing ships to steamships, 1850-1890: a study in technological change and its diffusion. In: McCloskey, D. N. eds. *Essays on a Mature Economy: Britain after 1840*. London: Methuen & Co. Ltd.

Harley, C. K. (1988) Ocean Freight Rates and Productivity, 1740-1913: The Primacy of Mechanical Invention Reaffirmed. *The Journal of Economic History*, 48 (4): 851-876.

Heston, A., Summers, R. and Aten, B. (2006) *Penn World Table Version 6.2*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006. http://dc2.chass.utoronto.ca/pwt.

Hogan, W. T. (1971) Economic History of the Iron and Steel Industry in the United States. Lexington, MA.: D.C. Heath and Company.

Hogan, W. T. (1994) Steel in the 21st century: Competition forges a new world order. New York: Lexington Books.

Hogan Steel Archive (2008) Fordham University. http://www.library.fordham.edu/steelarchive/steelindex.html (accessed 17/9/2008).

Hollenstein, H. and Woerter, M. (2008) Inter- and intra-firm diffusion of technology: The example of E-commerce: An analysis based on Swiss firm-level data. *Research Policy*, 37 (3): 545-564.

Hoover, K. D. and Perez, S. J. (2004) Truth and Robustness in Cross-Country Growth Regressions. *Oxford Bulletin of Economics and Statistics*, 66 (5): 765-798.

Hoppe, H. C. (2002) The Timing of New Technology Adoption: Theoretical Models and Empirical Evidence. *Manchester School*, 70 (1): 56-76.

Howells, J. (2002) The Response of Old Technology Incumbents to Technological Competition - Does the Sailing Ship Effect Exist? *Journal of Management Studies*, 39 (7): 887-906.

Hsiao, C. (1986) *Analysis of Panel Data*. Cambridge: Cambridge University Press.

Huntington, S. P. (1968) *Political order in changing societies*. New Haven and London: Yale University Press.

International Monetary Fund (2008) *International Financial Statistics*. September 2008 edition. Accessed via ESDS International at Mimas, University of Manchester. http://esds.mcc.ac.uk/wds_ifs, accessed September 2008.

Ireland, N. and Stoneman, P. (1985) Order effects, perfect foresight and intertemporal price discrimination. *Recherches Economiques de Louvain*, 51 (1): 7-20.

Ireland, N. and Stoneman, P. (1986) Technological Diffusion, Expectations and Welfare. *Oxford Economic Papers*, 38 (2): 283-304.

Judson, R. A. and Owen, A. L. (1999) Estimating dynamic panel data models: a guide for macroeconomists. *Economics Letters*, 65 (1): 9-15.

Karshenas, M. and Stoneman, P. (1993) Rank, stock, order and epidemic effects in the diffusion of new technologies: An empirical model. *RAND Journal of Economics*, 24 (4): 503-528.

Keller, W. (2004) International Technology Diffusion. *Journal of Economic Literature*, 42 (3): 752-782.

Kiiski, S. and Pohjola, M. (2002) Cross-country diffusion of the Internet. *Information Economics and Policy*, 14 (2): 297-310.

Kipping, M., Ranieri, R. and Dankers, J. (2001) The Emergence of New Competitor Nations in the European Steel Industry: Italy and the Netherlands, 1945-65. *Business History*, 43 (1): 69-96.

Knauerhase, R. (1968) The Compound Steam Engine and Productivity Changes in the German Merchant Marine Fleet, 1871-1887. *The Journal of Economic History*, 28 (3): 390-403.

Knudson, M. K. (1991) Incorporating Technological Change in Diffusion Models. *American Journal of Agricultural Economics*, 73 (3): 724-733.

Kumar, V. and Krishnan, T. V. (2002) Multinational Diffusion Models: An Alternative Framework. *Marketing Science*, 21 (3): 318-330.

Levine, R. and Renelt, D. (1992) A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review* 82 (4): 942-963.

MacKinnon, J. G. (1991) Critical Values for Cointegration Tests. In: Engle, R. F. and Granger, C. W. J. eds. *Long-Run Economic Relationships: Readings in Cointegration*. Oxford: Oxford University Press.

Maddala, G. S. and Knight, P. T. (1967) International Diffusion of Technical Change - A Case Study of the Oxygen Steel Making Process. *The Economic Journal*, 77 (307): 531-558.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992) A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107 (2): 407-437.

Mansfield, E. (1961) Technical Change and the Rate of Imitation. *Econometrica*, 29 (4): 741-766.

Mansfield, E. (1968) Industrial Research and Technological Innovation. New York: Norton.

McAdams, A. K. (1967) Big Steel, Invention, and Innovation Reconsidered. *The Quarterly Journal of Economics*, 81 (3): 457-474.

Meyer, J. R. and Herregat, G. (1974) The basic oxygen steel process. In: Nabseth, L. and Ray, G. F. eds. *The diffusion of new industrial processes: An international study*. London and New York: Cambridge University Press.

Mitchell, B. R. (1998) *International Historical Statistics*. 4. ed. Basingstoke: Palgrave Macmillan.

Mitchell, B. R. (2003) *International Historical Statistics*. 5. ed. Basingstoke: Palgrave Macmillan.

Nelson, R. R. and Phelps, E. S. (1966) Investment in humans, technological diffusion, and economic growth. *American Economic Review* 56 (2): 69-75.

OECD (various issues) The Iron and Steel Industry in 1965 and trends in 1966. Paris: OECD.

OECD (2004) Handbook for Internationally Comparative Education Statistics – Concepts, standards, definitions and classifications. http://www.esds.ac.uk/international/support/user_guides/oecd/Educhandbook.pdf, accessed 24/8/2009.

O'Malley, G.B. (2000) The Mineral Industries. In: *Technology in Australia 1788-1988*. Melbourne: Australian Academy of Technological Sciences and Engineering. Online edition published by Australian Science and Technology Heritage Centre 2000, updated 21/11/2001. http://www.austehc.unimelb.edu.au/tia/723.html, accessed 15/9/2009.

Oster, S. (1982) The Diffusion of Innovation among Steel Firms: The Basic Oxygen Furnace. *The Bell Journal of Economics*, 13 (1): 45-56.

Perkins, R. and Neumayer, E. (2005) The international diffusion of new technologies: A multitechnology analysis of latecomer advantage and global economic integration. *Annals of the Association of American Geographers*, 95 (4): 789-808.

Perron, P. (1989) The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57 (6): 1361-1401.

Przeworski, A., Alvarez, M. E., Cheibub, J. A. and Limongi, F. (2000) *Democracy and Development – Political Institutions and Well-Being in the World, 1950-1990*. Cambridge etc.: Cambridge University Press.

Reinganum, J. F. (1981a) On the Diffusion of New Technology: A Game Theoretic Approach. *The Review of Economic Studies*, 48 (3): 395-405.

Reinganum, J. F. (1981b) Market structure and the diffusion of new technology. *Bell Journal of Economics*, 12 (2): 618-624.

Richards, P.N. (2004) Ian Munro McLennan 1909-1998. *Historical Records of Australian Science* 15(2). http://www.science.org.au/academy/memoirs/mclennan.htm, accessed 15/9/2009.

Rogers, E. M. (1995) *Diffusion of Innovations*. 4th ed. New York: Free Press.

Rosegger, G. (1980) Exploratory evaluations of major technological innovations: Basic oxygen furnaces and continuous casting. In: Gold, B., Rosegger, G. and Boylan, M. G. eds. *Evaluating Technological Innovations: Methods, Expectations, and Findings*. Lexington, MA. and Toronto: Lexington Books.

Sala-i-Martin, X. (1997) I Just Ran Two Million Regressions. *American Economic Review*, 87 (2): 178-183.

Saloner, G. and Shepard, A. (1995) Adoption of technologies with network effects: An empirical examination fo the adoption of automatic teller machines. *Rand Journal of Economics*, 26 (3): 479-501.

Samstag, T. and Joshi, R. (2005) *Norwegian shipping - the past, present and the future*. Oslo: Horn Forlag AS.

Schumpeter, J. A. (1934) *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Singer, Jacques (1969) *Trade liberalization and the Canadian steel industry*. Toronto: University of Toronto Press.

Skinner, J. S. and Staigler, D. (2005) Technology adoption from hybrid corn to beta blockers. *NBER working paper series* No. 11251. Cambridge, MA: National Bureau of Economic Research, April 2005.

Stoneman, P. (2002) *The Economics of Technological Diffusion*. Oxford and Malden, MA: Blackwell.

Stoneman, P. and Battisti, G. (2010, forthcoming) The diffusion of new technology. In: Hall, B. H. and Rosenberg, N. eds. *Handbook of the Economics of Technical Change*. North Holland: Elsevier.

Stoneman, P. and Ireland, N. J. (1983) The role of supply factors in the diffusion of new process technology. *Economic Journal*, 93 (Supplement: Conference Papers (1983)): 66-78.

Stoneman, P. and Kwon, M. J. (1996) Technology adoption and firm profitability. *Economic Journal*, 106 (437): 952-962.

Stoneman, P. and Toivanen, O. (2006) Technological diffusion under uncertainty: a real options model applied to the comparative international diffusion of robot technology. In: Antonelli, C., Foray, D., Hall, B. H. and Steinmueller, W. E. eds.

New Frontiers in the Economics of Innovation and New Technology - Essays in honour of Paul A. David. Cheltenham and Northampton, MA.: Edward Elgar.

Strang, D. and Soule, S. A. (1998) Diffusion in organizations and social movements: From Hybrid Corn to Poison Pills. *Annual Review of Sociology*, 24: 265-290.

Sumrall, J. B., Jr. (1982) Diffusion of the Basic Oxygen Furnace in the U.S. Steel Industry. *Journal of Industrial Economics*, 30 (4): 421-437.

Tarr, D. G. (1985) The Efficient Diffusion of Steel Technology Across Nations. *Journal of Public Policy*, 5 (4): 541-560.

Teece, D. J. (1977) Technology transfer by multinational firms: the resource cost of transferring technological know-how. *Economic Journal*, 87 (346): 242-261.

Teece, D. J., Pisano, G. and Shuen, A. (1997) Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18 (7): 509-533.

Tellis, G. J., Stremersch, S. and Yin, E. (2003) The International Takeoff of New Products: The Role of Economics, Culture, and Country Innovativeness. *Marketing Science*, 22 (2): 188-208.

United Nations Economic Commission for Europe (various issues) *The European Steel Market in 1967*. Geneva: United Nations.

Verbeek, M. (2000) *A Guide to Modern Econometrics*. Chichester: John Wiley & Sons Ltd.